AF

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 October 2003 (09.10.2003)

PCT

(10) International Publication Number
WO 03/083085 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US03/09797

(22) International Filing Date: 27 March 2003 (27.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/368,840 28 March 2002 (28.03.2002) US
60/375,637 26 April 2002 (26.04.2002) US

(71) Applicant *(for all designated States except US)*: **INCYTE CORPORATION** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MARQUIS, Joseph, P.** [US/US]; 4428 Lazy Lane, San Jose, CA 95135 (US). **LEE, Soo, Y.** [KR/US]; 40 Westdale Avenue, Daly City, CA 94015 (US). **EMERLING, Brooke, M.** [US/US]; 833 W. Buena Avenue #2108, Chicago, IL 60613 (US). **HAFALIA, April, J.A.** [US/US]; 15 Midvale Drive, Daly City, CA 94015 (US). **KHARE, Reena** [IN/US]; 12650 Orella Court, Saratoga, CA 95070 (US). **KABLE, Amy, E.** [US/US]; 2345 Polk Street #4, San Francisco, CA 94109 (US). **RICHARDSON, Thomas, W.** [US/US]; 616 Canyon Road #107, Redwood City, CA 94062 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue #5D, San Francisco, CA 94122 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square, #712, Union City, CA 94587 (US). **BECHA, Shanya, D.** [US/US]; 1 Saint Francis #5508, San Francisco, CA 94107 (US). **MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95014 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **TRAN, Uyen, K.** [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US). **ISON, Craig, H.** [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). **LINDQUIST, Erika, A.** [US/US]; 2394 Mariner Square Drive #C-121, Alameda, CA 94501 (US). **JIANG, Xin** [US/US]; 14371 Elva Avenue, Saratoga, CA 95070 (US). **JACKSON, Alan, A.** [US/US]; 1541 Elwood Drive, Los Gatos, CA 95032 (US). **WILSON, Amy, D.** [US/US]; 1056 Continentals Way, #27, Belmont, CA 94002 (US). **JIN, Pei** [US/US]; 320 Curtner Avenue #D, Palo Alto, CA 94306 (US). **CHANG, Hsin-Ru** [US/US]; 326 Treasure Island Drive, Belmont, CA 94002 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Inycte Corporation, 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *without international search report and to be republished upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 03/083085 A2

(54) Title: TRANSPORTERS AND ION CHANNELS

(57) Abstract: Various embodiments of the invention provide human transporters and ion channels (TRICH) and polynucleotides which identify and encode TRICH. Embodiments of the invention also provide expression vectors, host cells, antibodies, agonists, and antagonists. Other embodiments provide methods for diagnosing, treating, or preventing disorders associated with aberrant expression of TRICH.

# TRANSPORTERS AND ION CHANNELS

## TECHNICAL FIELD

The invention relates to novel nucleic acids, transporters and ion channels encoded by these nucleic acids, and to the use of these nucleic acids and proteins in the diagnosis, treatment, and prevention of transport, neurological, muscle, immunological and cell proliferative disorders. The invention also relates to the assessment of the effects of exogenous compounds on the expression of nucleic acids and transporters and ion channels.

## BACKGROUND OF THE INVENTION

Eukaryotic cells are surrounded and subdivided into functionally distinct organelles by hydrophobic lipid bilayer membranes which are highly impermeable to most polar molecules. Cells and organelles require transport proteins to import and export essential nutrients and metal ions including $K^+$, $NH_4^+$, $P_i$, $SO_4^{2-}$, sugars, and vitamins, as well as various metabolic waste products. Transport proteins also play roles in antibiotic resistance, toxin secretion, ion balance, synaptic neurotransmission, kidney function, intestinal absorption, tumor growth, and other diverse cell functions (Griffith, J. and C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, pp. 3-29). Transport can occur by a passive concentration-dependent mechanism, or can be linked to an energy source such as ATP hydrolysis or an ion gradient. Proteins that function in transport include carrier proteins, which bind to a specific solute and undergo a conformational change that translocates the bound solute across the membrane, and channel proteins, which form hydrophilic pores that allow specific solutes to diffuse through the membrane down an electrochemical solute gradient.

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelium contains a variety of symporter systems driven by the sodium gradient that exists across the plasma membrane. Sodium moves into the cell down its electrochemical gradient and brings the solute into the cell with it. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous $Na^+/K^+$ ATPase system. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). All three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini. NIS plays a crucial role in the evaluation, diagnosis, and treatment of

various thyroid pathologies because it is the molecular basis for radioiodide thyroid-imaging techniques and for specific targeting of radioisotopes to the thyroid gland (Levy, O. et al. (1997) Proc. Natl. Acad. Sci. USA 94:5568-5573). SMVT is expressed in the intestinal mucosa, kidney, and placenta, and is implicated in the transport of the water-soluble vitamins, e.g., biotin and pantothenate (Prasad, P.D. et al. (1998) J. Biol. Chem. 273:7501-7506).

One of the largest families of transporters is the major facilitator superfamily (MFS), also called the uniporter-symporter-antiporter family. MFS transporters are single polypeptide carriers that transport small solutes in response to ion gradients. Members of the MFS are found in all classes of living organisms, and include transporters for sugars, oligosaccharides, phosphates, nitrates, nucleosides, monocarboxylates, and drugs. MFS transporters found in eukaryotes all have a structure comprising 12 transmembrane segments (Pao, S.S. et al. (1998) Microbiol. Molec. Biol. Rev. 62:1-34). The largest family of MFS transporters is the sugar transporter family, which includes the seven glucose transporters (GLUT1-GLUT7) found in humans that are required for the transport of glucose and other hexose sugars. These glucose transport proteins have unique tissue distributions and physiological functions. GLUT1 provides many cell types with their basal glucose requirements and transports glucose across epithelial and endothelial barrier tissues; GLUT2 facilitates glucose uptake or efflux from the liver; GLUT3 regulates glucose supply to neurons; GLUT4 is responsible for insulin-regulated glucose disposal; and GLUT5 regulates fructose uptake into skeletal muscle. Defects in glucose transporters are involved in a recently identified neurological syndrome causing infantile seizures and developmental delay, as well as glycogen storage disease, Fanconi-Bickel syndrome, and non-insulin-dependent diabetes mellitus (Mueckler, M. (1994) Eur. J. Biochem. 219:713-725; Longo, N. and L.J. Elsas (1998) Adv. Pediatr. 45:293-313).

Monocarboxylate anion transporters are proton-coupled symporters with a broad substrate specificity that includes L-lactate, pyruvate, and the ketone bodies acetate, acetoacetate, and beta-hydroxybutyrate. At least seven isoforms have been identified to date. The isoforms are predicted to have twelve transmembrane (TM) helical domains with a large intracellular loop between TM6 and TM7, and play a critical role in maintaining intracellular pH by removing the protons that are produced stoichiometrically with lactate during glycolysis. The best characterized $H^+$-monocarboxylate transporter is that of the erythrocyte membrane, which transports L-lactate and a wide range of other aliphatic monocarboxylates. Other cells possess $H^+$-linked monocarboxylate transporters with differing substrate and inhibitor selectivities. In particular, cardiac muscle and tumor cells have transporters that differ in their $K_m$ values for certain substrates, including stereoselectivity for L- over D-lactate, and in their sensitivity to inhibitors. There are $Na^+$-monocarboxylate

cotransporters on the luminal surface of intestinal and kidney epithelia, which allow the uptake of lactate, pyruvate, and ketone bodies in these tissues. In addition, there are specific and selective transporters for organic cations and organic anions in organs including the kidney, intestine and liver. Organic anion transporters are selective for hydrophobic, charged molecules with electron-attracting side groups. Organic cation transporters, such as the ammonium transporter, mediate the secretion of a variety of drugs and endogenous metabolites, and contribute to the maintenance of intercellular pH (Poole, R.C. and A.P. Halestrap (1993) Am. J. Physiol. 264:C761-C782; Price, N.T. et al. (1998) Biochem. J. 329:321-328; and Martinelle, K. and I. Haggstrom (1993) J. Biotechnol. 30:339-350).

ATP-binding cassette (ABC) transporters are members of a superfamily of membrane proteins that transport substances ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. ABC transporters consist of four modules: two nucleotide-binding domains (NBD), which hydrolyze ATP to supply the energy required for transport, and two membrane-spanning domains (MSD), each containing six putative transmembrane segments. These four modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product contains a single NBD and MSD. These "half-molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system. Several genetic diseases are attributed to defects in ABC transporters, such as the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel), adrenoleukodystrophy (adrenoleukodystrophy protein, ALDP), Zellweger syndrome (peroxisomal membrane protein-70, PMP70), and hyperinsulinemic hypoglycemia (sulfonylurea receptor, SUR). Overexpression of the multidrug resistance (MDR) protein, another ABC transporter, in human cancer cells makes the cells resistant to a variety of cytotoxic drugs used in chemotherapy (Taglicht, D. and S. Michaelis (1998) Meth. Enzymol. 292:130-162).

A number of metal ions such as iron, zinc, copper, cobalt, manganese, molybdenum, selenium, nickel, and chromium are important as cofactors for a number of enzymes. For example, copper is involved in hemoglobin synthesis, connective tissue metabolism, and bone development, by acting as a cofactor in oxidoreductases such as superoxide dismutase, ferroxidase (ceruloplasmin), and lysyl oxidase. Copper and other metal ions must be provided in the diet, and are absorbed by transporters in the gastrointestinal tract. Plasma proteins transport the metal ions to the liver and other target organs, where specific transporters move the ions into cells and cellular organelles as needed. Imbalances in metal ion metabolism have been associated with a number of disease states (Danks, D.M. (1986) J. Med. Genet. 23:99-106).

Transport of fatty acids across the plasma membrane can occur by diffusion, a high capacity, low affinity process. However, under normal physiological conditions a significant fraction of fatty acid transport appears to occur via a high affinity, low capacity protein-mediated transport process. Fatty acid transport protein (FATP), an integral membrane protein with four transmembrane segments, is expressed in tissues exhibiting high levels of plasma membrane fatty acid flux, such as muscle, heart, and adipose. Expression of FATP is upregulated in 3T3-L1 cells during adipose conversion, and expression in COS7 fibroblasts elevates uptake of long-chain fatty acids (Hui, T.Y. et al. (1998) J. Biol. Chem. 273:27420-27429).

The lipocalin superfamily constitutes a phylogenetically conserved group of more than forty proteins that function as extracellular ligand-binding proteins which bind and transport small hydrophobic molecules. Members of this family function as carriers of retinoids, odorants, chromophores, pheromones, allergens, and sterols, and in a variety of processes including nutrient transport, cell growth regulation, immune response, and prostaglandin synthesis. A subset of these proteins may be multifunctional, serving as either a biosynthetic enzyme or as a specific enzyme inhibitor. (Tanaka, T. et al. (1997) J. Biol. Chem. 272:15789-15795; and van't Hof, W. et al. (1997) J. Biol. Chem. 272:1837-1841.)

Members of the lipocalin family display unusually low levels of overall sequence conservation. Pairwise sequence identity often falls below 20%. Sequence similarity between family members is limited to conserved cysteines which form disulfide bonds and three motifs which form a juxtaposed cluster that functions as a target cell recognition site. The lipocalins share an eight stranded, anti-parallel beta-sheet which folds back on itself to form a continuously hydrogen-bonded beta-barrel. The pocket formed by the barrel functions as an internal ligand binding site. Seven loops (L1 to L7) form short beta-hairpins, except loop L1 which is a large omega loop that forms a lid to partially close the internal ligand-binding site (Flower (1996) Biochem. J. 318:1-14).

Lipocalins are important transport molecules. Each lipocalin associates with a particular ligand and delivers that ligand to appropriate target sites within the organism. Retinol-binding protein (RBP), one of the best characterized lipocalins, transports retinol from stores within the liver to target tissues. Apolipoprotein D (apo D), a component of high density lipoproteins (HDLs) and low density lipoproteins (LDLs), functions in the targeted collection and delivery of cholesterol throughout the body. Lipocalins are also involved in cell regulatory processes. Apo D, which is identical to gross-cystic-disease-fluid protein (GCDFP)-24, is a progesterone/pregnenolone-binding protein expressed at high levels in breast cyst fluid. Secretion of apo D in certain human breast cancer cell lines is accompanied by reduced cell proliferation and progression of cells to a more differentiated phenotype.

Similarly, apo D and another lipocalin, $\alpha_1$-acid glycoprotein (AGP), are involved in nerve cell regeneration. AGP is also involved in anti-inflammatory and immunosuppressive activities. AGP is one of the positive acute-phase proteins (APP); circulating levels of AGP increase in response to stress and inflammatory stimulation. AGP accumulates at sites of inflammation where it inhibits platelet and neutrophil activation and inhibits phagocytosis. The immunomodulatory properties of AGP are due to glycosylation. AGP is 40% carbohydrate, making it unusually acidic and soluble. The glycosylation pattern of AGP changes during acute-phase response, and deglycosylated AGP has no immunosuppressive activity (Flower (1994) FEBS Lett. 354:7-11; Flower (1996) *supra*).

The lipocalin superfamily also includes several animal allergens, including the mouse major urinary protein (mMUP), the rat α-2-microgloobulin (rA2U), the bovine β-lactoglobulin (βlg), the cockroach allergen (Bla g4), bovine dander allergen (Bos d2), and the major horse allergen, designated *Equus caballus* allergen 1 (Equ c1). Equ c1 is a powerful allergen responsible for about 80% of anti-horse IgE antibody response in patients who are chronically exposed to horse allergens. It appears that lipocalins may contain a common structure that is able to induce the IgE response (Gregoire, C. et al., (1996) J. Biol. Chem. 271:32951-32959).

Lipocalins are used as diagnostic and prognostic markers in a variety of disease states. The plasma level of AGP is monitored during pregnancy and in diagnosis and prognosis of conditions including cancer chemotherapy, renal disfunction, myocardial infarction, arthritis, and multiple sclerosis. RBP is used clinically as a marker of tubular reabsorption in the kidney, and apo D is a marker in gross cystic breast disease (Flower (1996) *supra*). Additionally, the use of lipocalin animal allergens may help in the diagnosis of allergic reactions to horses (Gregoire *supra*), pigs, cockroaches, mice and rats.

Mitochondrial carrier proteins are transmembrane-spanning proteins which transport ions and charged metabolites between the cytosol and the mitochondrial matrix. Examples include the ADP, ATP carrier protein; the 2-oxoglutarate/malate carrier; the phosphate carrier protein; the pyruvate carrier; the dicarboxylate carrier which transports malate, succinate, fumarate, and phosphate; the tricarboxylate carrier which transports citrate and malate; and the Grave's disease carrier protein, a protein recognized by IgG in patients with active Grave's disease, an autoimmune disorder resulting in hyperthyroidism. Proteins in this family consist of three tandem repeats of an approximately 100 amino acid domain, each of which contains two transmembrane regions (Stryer, L. (1995) Biochemistry, W.H. Freeman and Company, New York NY, p. 551; PROSITE PDOC00189 Mitochondrial energy transfer proteins signature; Online Mendelian Inheritance in Man (OMIM) *275000 Graves Disease).

This class of transporters also includes the mitochondrial uncoupling proteins, which create proton leaks across the inner mitochondrial membrane, thus uncoupling oxidative phosphorylation from ATP synthesis. The result is energy dissipation in the form of heat. Mitochondrial uncoupling proteins have been implicated as modulators of thermoregulation and metabolic rate, and have been proposed as potential targets for drugs against metabolic diseases such as obesity (Ricquier, D. et al. (1999) J. Int. Med. 245:637-642).

**Ion Channels**

The electrical potential of a cell is generated and maintained by controlling the movement of ions across the plasma membrane. The movement of ions requires ion channels, which form ion-selective pores within the membrane. There are two basic types of ion channels, ion transporters and gated ion channels. Ion transporters utilize the energy obtained from ATP hydrolysis to actively transport an ion against the ion's concentration gradient. Gated ion channels allow passive flow of an ion down the ion's electrochemical gradient under restricted conditions. Together, these types of ion channels generate, maintain, and utilize an electrochemical gradient that is used in 1) electrical impulse conduction down the axon of a nerve cell, 2) transport of molecules into cells against concentration gradients, 3) initiation of muscle contraction, and 4) endocrine cell secretion.

<u>Ion Transporters</u>

Ion transporters generate and maintain the resting electrical potential of a cell. Utilizing the energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient. These transmembrane ATPases are divided into three families. The phosphorylated (P) class ion transporters, including $Na^+$-$K^+$ ATPase, $Ca^{2+}$-ATPase, and $H^+$-ATPase, are activated by a phosphorylation event. P-class ion transporters are responsible for maintaining resting potential distributions such that cytosolic concentrations of $Na^+$ and $Ca^{2+}$ are low and cytosolic concentration of $K^+$ is high. The vacuolar (V) class of ion transporters includes $H^+$ pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of $H^+$ pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate ($P_i$).

The P-ATPases are hexamers of a 100 kD subunit with ten transmembrane domains and several large cytoplasmic regions that may play a role in ion binding (Scarborough, G.A. (1999) Curr. Opin. Cell Biol. 11:517-522). The V-ATPases are composed of two functional domains: the $V_1$ domain, a peripheral complex responsible for ATP hydrolysis; and the $V_0$ domain, an integral complex responsible for proton translocation across the membrane. The F-ATPases are structurally and

evolutionarily related to the V-ATPases. The F-ATPase $F_0$ domain contains 12 copies of the c subunit, a highly hydrophobic protein composed of two transmembrane domains and containing a single buried carboxyl group in TM2 that is essential for proton transport. The V-ATPase $V_0$ domain contains three types of homologous c subunits with four or five transmembrane domains and the essential carboxyl group in TM4 or TM3. Both types of complex also contain a single a subunit that may be involved in regulating the pH dependence of activity (Forgac, M. (1999) J. Biol. Chem. 274:12951-12954).

The resting potential of the cell is utilized in many processes involving carrier proteins and gated ion channels. Carrier proteins utilize the resting potential to transport molecules into and out of the cell. Amino acid and glucose transport into many cells is linked to sodium ion co-transport (symport) so that the movement of $Na^+$ down an electrochemical gradient drives transport of the other molecule up a concentration gradient. Similarly, cardiac muscle links transfer of $Ca^{2+}$ out of the cell with transport of $Na^+$ into the cell (antiport).

<u>Gated Ion Channels</u>

Gated ion channels control ion flow by regulating the opening and closing of pores. The ability to control ion flux through various gating mechanisms allows ion channels to mediate such diverse signaling and homeostatic functions as neuronal and endocrine signaling, muscle contraction, fertilization, and regulation of ion and pH balance. Gated ion channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open their pores in response to mechanical stress; voltage-gated channels (e.g., $Na^+$, $K^+$, $Ca^{2+}$, and $Cl^-$ channels) open their pores in response to changes in membrane potential; and ligand-gated channels (e.g., acetylcholine-, serotonin-, and glutamate-gated cation channels, and GABA- and glycine-gated chloride channels) open their pores in the presence of a specific ion, nucleotide, or neurotransmitter. The gating properties of a particular ion channel (i.e., its threshold for and duration of opening and closing) are sometimes modulated by association with auxiliary channel proteins and/or post translational modifications, such as phosphorylation.

Mechanically-gated or mechanosensitive ion channels act as transducers for the senses of touch, hearing, and balance, and also play important roles in cell volume regulation, smooth muscle contraction, and cardiac rhythm generation. A stretch-inactivated channel (SIC) was recently cloned from rat kidney. The SIC channel belongs to a group of channels which are activated by pressure or stress on the cell membrane and conduct both $Ca^{2+}$ and $Na^+$ (Suzuki, M. et al. (1999) J. Biol. Chem. 274:6330-6335).

The pore-forming subunits of the voltage-gated cation channels form a superfamily of ion

channel proteins. The characteristic domain of these channel proteins comprises six transmembrane domains (S1-S6), a pore-forming region (P) located between S5 and S6, and intracellular amino and carboxy termini. In the $Na^+$ and $Ca^{2+}$ subfamilies, this domain is repeated four times, while in the $K^+$ channel subfamily, each channel is formed from a tetramer of either identical or dissimilar subunits. The P region contains information specifying the ion selectivity for the channel. In the case of $K^+$ channels, a GYG tripeptide is involved in this selectivity (Ishii, T.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11651-11656).

Voltage-gated $Na^+$ and $K^+$ channels are necessary for the function of electrically excitable cells, such as nerve and muscle cells. Action potentials, which lead to neurotransmitter release and muscle contraction, arise from large, transient changes in the permeability of the membrane to $Na^+$ and $K^+$ ions. Depolarization of the membrane beyond the threshold level opens voltage-gated $Na^+$ channels. Sodium ions flow into the cell, further depolarizing the membrane and opening more voltage-gated $Na^+$ channels, which propagates the depolarization down the length of the cell. Depolarization also opens voltage-gated potassium channels. Consequently, potassium ions flow outward, which leads to repolarization of the membrane. Voltage-gated channels utilize charged residues in the fourth transmembrane segment (S4) to sense voltage change. The open state lasts only about 1 millisecond, at which time the channel spontaneously converts into an inactive state that cannot be opened irrespective of the membrane potential. Inactivation is mediated by the channel's N-terminus, which acts as a plug that closes the pore. The transition from an inactive to a closed state requires a return to resting potential.

Voltage-gated $Na^+$ channels are heterotrimeric complexes composed of a 260 kDa pore-forming $\alpha$ subunit that associates with two smaller auxiliary subunits, $\beta1$ and $\beta2$. The $\beta2$ subunit is a integral membrane glycoprotein that contains an extracellular Ig domain, and its association with $\alpha$ and $\beta1$ subunits correlates with increased functional expression of the channel, a change in its gating properties, as well as an increase in whole cell capacitance due to an increase in membrane surface area (Isom, L.L. et al. (1995) Cell 83:433-442).

Non voltage-gated $Na^+$ channels include the members of the amiloride-sensitive $Na^+$ channel/degenerin (NaC/DEG) family. Channel subunits of this family are thought to consist of two transmembrane domains flanking a long extracellular loop, with the amino and carboxyl termini located within the cell. The NaC/DEG family includes the epithelial $Na^+$ channel (ENaC) involved in $Na^+$ reabsorption in epithelia including the airway, distal colon, cortical collecting duct of the kidney, and exocrine duct glands. Mutations in ENaC result in pseudohypoaldosteronism type 1 and Liddle's syndrome (pseudohyperaldosteronism). The NaC/DEG family also includes the recently characterized

$H^+$-gated cation channels or acid-sensing ion channels (ASIC). ASIC subunits are expressed in the brain and form heteromultimeric $Na^+$-permeable channels. These channels require acid pH fluctuations for activation. ASIC subunits show homology to the degenerins, a family of mechanically-gated channels originally isolated from C. elegans. Mutations in the degenerins cause neurodegeneration. ASIC subunits may also have a role in neuronal function, or in pain perception, since tissue acidosis causes pain (Waldmann, R. and M. Lazdunski (1998) Curr. Opin. Neurobiol. 8:418-424; Eglen, R.M. et al. (1999) Trends Pharmacol. Sci. 20:337-342).

$K^+$ channels are located in all cell types, and may be regulated by voltage, ATP concentration, or second messengers such as $Ca^{2+}$ and cAMP. In non-excitable tissue, $K^+$ channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes, $K^+$ channels are responsible for setting the resting membrane potential. The cytosol contains non-diffusible anions and, to balance this net negative charge, the cell contains a $Na^+$-$K^+$ pump and ion channels that provide the redistribution of $Na^+$, $K^+$, and $Cl^-$. The pump actively transports $Na^+$ out of the cell and $K^+$ into the cell in a 3:2 ratio. Ion channels in the plasma membrane allow $K^+$ and $Cl^-$ to flow by passive diffusion. Because of the high negative charge within the cytosol, $Cl^-$ flows out of the cell. The flow of $K^+$ is balanced by an electromotive force pulling $K^+$ into the cell, and a $K^+$ concentration gradient pushing $K^+$ out of the cell. Thus, the resting membrane potential is primarily regulated by $K^+$ flow (Salkoff, L. and T. Jegla (1995) Neuron 15:489-492).

Potassium channel subunits of the *Shaker*-like superfamily all have the characteristic six transmembrane/1 pore domain structure. Four subunits combine as homo- or heterotetramers to form functional K channels. These pore-forming subunits also associate with various cytoplasmic $\beta$ subunits that alter channel inactivation kinetics. The *Shaker*-like channel family includes the voltage-gated $K^+$ channels as well as the delayed rectifier type channels such as the human ether-a-go-go related gene (HERG) associated with long QT, a cardiac dysrythmia syndrome (Curran, M.E. (1998) Curr. Opin. Biotechnol. 9:565-572; Kaczorowski, G.J. and M.L. Garcia (1999) Curr. Opin. Chem. Biol. 3:448-458).

A second superfamily of $K^+$ channels is composed of the inward rectifying channels (Kir). Kir channels have the property of preferentially conducting $K^+$ currents in the inward direction. These proteins consist of a single potassium selective pore domain and two transmembrane domains, which correspond to the fifth and sixth transmembrane domains of voltage-gated $K^+$ channels. Kir subunits also associate as tetramers. The Kir family includes ROMK1, mutations in which lead to Bartter syndrome, a renal tubular disorder. Kir channels are also involved in regulation of cardiac pacemaker

activity, seizures and epilepsy, and insulin regulation (Doupnik, C.A. et al. (1995) Curr. Opin. Neurobiol. 5:268-277; Curran, *supra*).

The recently recognized TWIK $K^+$ channel family includes the mammalian TWIK-1, TREK-1 and TASK proteins. Members of this family possess an overall structure with four transmembrane domains and two P domains. These proteins are probably involved in controlling the resting potential in a large set of cell types (Duprat, F. et al. (1997) EMBO J 16:5464-5471).

The voltage-gated $Ca^{2+}$ channels have been classified into several subtypes based upon their electrophysiological and pharmacological characteristics. L-type $Ca^{2+}$ channels are predominantly expressed in heart and skeletal muscle where they play an essential role in excitation-contraction coupling. T-type channels are important for cardiac pacemaker activity, while N-type and P/Q-type channels are involved in the control of neurotransmitter release in the central and peripheral nervous system. The L-type and N-type voltage-gated $Ca^{2+}$ channels have been purified and, though their functions differ dramatically, they have similar subunit compositions. The channels are composed of three subunits. The $\alpha_1$ subunit forms the membrane pore and voltage sensor, while the $\alpha_2\delta$ and $\beta$ subunits modulate the voltage-dependence, gating properties, and the current amplitude of the channel. These subunits are encoded by at least six $\alpha_1$, one $\alpha_2\delta$, and four $\beta$ genes. A fourth subunit, $\gamma$, has been identified in skeletal muscle (Walker, D. et al. (1998) J. Biol. Chem. 273:2361-2367; McCleskey, E.W. (1994) Curr. Opin. Neurobiol. 4:304-312).

The high-voltage-activated $Ca^{2+}$ channels that have been characterized biochemically include complexes of a pore-forming alpha1 subunit of approximately 190-250 kDa; a transmembrane complex of alpha2 and delta subunits; an intracellular beta subunit; and in some cases a transmembrane gamma subunit. A variety of alpha1 subunits, alpha2delta complexes, beta subunits, and gamma subunits are known. The Cav1 family of alpha1 subunits conduct L-type $Ca^{2+}$ currents, which initiate muscle contraction, endocrine secretion, and gene transcription, and are regulated primarily by second messenger-activated protein phosphorylation pathways. The Cav2 family of alpha1 subunits conduct N-type, P/Q-type, and R-type $Ca^{2+}$ currents, which initiate rapid synaptic transmission and are regulated primarily by direct interaction with G proteins and SNARE proteins and secondarily by protein phosphorylation. The Cav3 family of alpha1 subunits conduct T-type $Ca^{2+}$ currents, which are activated and inactivated more rapidly and at more negative membrane potentials than other $Ca^{2+}$ current types. The distinct structures and patterns of regulation of these three families of $Ca^{2+}$ channels provide an array of $Ca^{2+}$ entry pathways in response to changes in membrane potential and a range of possibilities for regulation of $Ca^{2+}$ entry by second messenger pathways and interacting proteins (Catterall, W.A. (2000) Annu. Rev. Cell Dev. Biol. 16:521-555).

The alpha-2 subunit of the voltage-gated Ca $^{2+}$-channel may include one or more Cache domains. An extracellular Cache domain may be fused to an intracellular catalytic domain, such as the histidine kinase, PP2C phosphatase, GGDEF (a predicted diguanylate cyclase), HD-GYP (a predicted phosphodiesterase) or adenylyl cyclase domain, or to a noncatalytic domain, like the methyl-accepting, DNA-binding winged helix-turn-helix, GAF, PAS or HAMP (a domain found in istidine kinases, denylyl cyclases, ethyl-binding proteins and phosphatases). Small molecules are bound via the Cache domain and this signal is converted into diverse outputs depending on the intracellular domains (Anantharaman, V. and Aravind, L.(2000) Trends Biochem. Sci. 25:535-537).

The transient receptor family (Trp) of calcium ion channels are thought to mediate capacitative calcium entry (CCE). CCE is the $Ca^{2+}$ influx into cells to resupply $Ca^{2+}$ stores depleted by the action of inositol triphosphate (IP3) and other agents in response to numerous hormones and growth factors. Trp and Trp-like were first cloned from Drosophila and have similarity to voltage gated Ca $^{2+}$ channels in the S3 through S6 regions. This suggests that Trp and/or related proteins may form mammalian CCE channels (Zhu, X. et al. (1996) Cell 85:661-671; Boulay, G. et al. (1997) J. Biol. Chem. 272:29672-29680). Melastatin is a gene isolated in both the mouse and human, whose expression in melanoma cells is inversely correlated with melanoma aggressiveness *in vivo*. The human cDNA transcript corresponds to a 1533-amino acid protein having homology to members of the Trp family. It has been proposed that the combined use of malastatin mRNA expression status and tumor thickness might allow for the determination of subgroups of patients at both low and high risk for developing metastatic disease (Duncan, L.M. et al (2001) J. Clin. Oncol. 19:568-576).

Chloride channels are necessary in endocrine secretion and in regulation of cytosolic and organelle pH. In secretory epithelial cells, $Cl^-$ enters the cell across a basolateral membrane through an $Na^+$, $K^+/Cl^-$ cotransporter, accumulating in the cell above its electrochemical equilibrium concentration. Secretion of $Cl^-$ from the apical surface, in response to hormonal stimulation, leads to flow of $Na^+$ and water into the secretory lumen. The cystic fibrosis transmembrane conductance regulator (CFTR) is a chloride channel encoded by the gene for cystic fibrosis, a common fatal genetic disorder in humans. CFTR is a member of the ABC transporter family, and is composed of two domains each consisting of six transmembrane domains followed by a nucleotide-binding site. Loss of CFTR function decreases transepithelial water secretion and, as a result, the layers of mucus that coat the respiratory tree, pancreatic ducts, and intestine are dehydrated and difficult to clear. The resulting blockage of these sites leads to pancreatic insufficiency, "meconium ileus", and devastating "chronic obstructive pulmonary disease" (Al-Awqati, Q. et al. (1992) J. Exp. Biol. 172:245-266).

The voltage-gated chloride channels (CLC) are characterized by 10-12 transmembrane

domains, as well as two small globular domains known as CBS domains. The CLC subunits probably function as homotetramers. CLC proteins are involved in regulation of cell volume, membrane potential stabilization, signal transduction, and transepithelial transport. Mutations in CLC-1, expressed predominantly in skeletal muscle, are responsible for autosomal recessive generalized myotonia and autosomal dominant myotonia congenita, while mutations in the kidney channel CLC-5 lead to kidney stones (Jentsch, T.J. (1996) Curr. Opin. Neurobiol. 6:303-310).

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of $Na^+$ and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as $\gamma$-aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential. Neurotransmitter-gated ion channels have four transmembrane domains and probably function as pentamers (Jentsch, *supra*). Amino acids in the second transmembrane domain appear to be important in determining channel permeation and selectivity (Sather, W.A. et al. (1994) Curr. Opin. Neurobiol. 4:313-323).

Ligand-gated channels can be regulated by intracellular second messengers. For example, calcium-activated $K^+$ channels are gated by internal calcium ions. In nerve cells, an influx of calcium during depolarization opens $K^+$ channels to modulate the magnitude of the action potential (Ishi et al., *supra*). The large conductance (BK) channel has been purified from brain and its subunit composition determined. The $\alpha$ subunit of the BK channel has seven rather than six transmembrane domains in contrast to voltage-gated $K^+$ channels. The extra transmembrane domain is located at the subunit N-terminus. A 28-amino-acid stretch in the C-terminal region of the subunit (the "calcium bowl" region) contains many negatively charged residues and is thought to be the region responsible for calcium binding. The $\beta$ subunit consists of two transmembrane domains connected by a glycosylated extracellular loop, with intracellular N- and C-termini (Kaczorowski, *supra*; Vergara, C. et al. (1998) Curr. Opin. Neurobiol. 8:321-329).

Cyclic nucleotide-gated (CNG) channels are gated by cytosolic cyclic nucleotides. The best examples of these are the cAMP-gated $Na^+$ channels involved in olfaction and the cGMP-gated cation channels involved in vision. Both systems involve ligand-mediated activation of a G-protein coupled receptor which then alters the level of cyclic nucleotide within the cell. CNG channels also

represent a major pathway for $Ca^{2+}$ entry into neurons, and play roles in neuronal development and plasticity. CNG channels are tetramers containing at least two types of subunits, an $\alpha$ subunit which can form functional homomeric channels, and a $\beta$ subunit, which modulates the channel properties. All CNG subunits have six transmembrane domains and a pore forming region between the fifth and sixth transmembrane domains, similar to voltage-gated $K^+$ channels. A large C-terminal domain contains a cyclic nucleotide binding domain, while the N-terminal domain confers variation among channel subtypes (Zufall, F. et al. (1997) Curr. Opin. Neurobiol. 7:404-412).

The activity of other types of ion channel proteins may also be modulated by a variety of intracellular signaling proteins. Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Kir channels are activated by the binding of the $G\beta\gamma$ subunits of heterotrimeric G-proteins (Reimann, F. and F.M. Ashcroft (1999) Curr. Opin. Cell. Biol. 11:503-508). Other proteins are involved in the localization of ion channels to specific sites in the cell membrane. Such proteins include the PDZ domain proteins known as MAGUKs (membrane-associated guanylate kinases) which regulate the clustering of ion channels at neuronal synapses (Craven, S.E. and D.S. Bredt (1998) Cell 93:495-498).

**Disease Correlation**

The etiology of numerous human diseases and disorders can be attributed to defects in the transport of molecules across membranes. Defects in the trafficking of membrane-bound transporters and ion channels are associated with several disorders, e.g., cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, von Gierke disease, and certain forms of diabetes mellitus. Single-gene defect diseases resulting in an inability to transport small molecules across membranes include, e.g., cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease (van't Hoff, W.G. (1996) Exp. Nephrol. 4:253-262; Talente, G.M. et al. (1994) Ann. Intern. Med. 120:218-226; and Chillon, M. et al. (1995) New Engl. J. Med. 332:1475-1480).

Human diseases caused by mutations in ion channel genes include disorders of skeletal muscle, cardiac muscle, and the central nervous system. Mutations in the pore-forming subunits of sodium and chloride channels cause myotonia, a muscle disorder in which relaxation after voluntary contraction is delayed. Sodium channel myotonias have been treated with channel blockers. Mutations in muscle sodium and calcium channels cause forms of periodic paralysis, while mutations in the sarcoplasmic calcium release channel, T-tubule calcium channel, and muscle sodium channel cause malignant hyperthermia. Cardiac arrythmia disorders such as the long QT syndromes and idiopathic ventricular fibrillation are caused by mutations in potassium and sodium channels (Cooper,

E.C. and L.Y. Jan (1998) Proc. Natl. Acad. Sci. USA 96:4759-4766). All four known human idiopathic epilepsy genes code for ion channel proteins (Berkovic, S.F. and I.E. Scheffer (1999) Curr. Opin. Neurology 12:177-182). Other neurological disorders such as ataxias, hemiplegic migraine and hereditary deafness can also result from mutations in ion channel genes (Jen, J. (1999) Curr. Opin. Neurobiol. 9:274-280; Cooper, *supra*).

Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia. Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) Adv. Pharmacol. 39:47-98). Various classes of ion channels also play an important role in the perception of pain, and thus are potential targets for new analgesics. These include the vanilloid-gated ion channels, which are activated by the vanilloid capsaicin, as well as by noxious heat. Local anesthetics such as lidocaine and mexiletine which blockade voltage-gated $Na^+$ channels have been useful in the treatment of neuropathic pain (Eglen, *supra*).

Ion channels in the immune system have recently been suggested as targets for immunomodulation. T-cell activation depends upon calcium signaling, and a diverse set of T-cell specific ion channels has been characterized that affect this signaling process. Channel blocking agents can inhibit secretion of lymphokines, cell proliferation, and killing of target cells. A peptide antagonist of the T-cell potassium channel Kv1.3 was found to suppress delayed-type hypersensitivity and allogenic responses in pigs, validating the idea of channel blockers as safe and efficacious immunosuppressants (Cahalan, M.D. and K.G. Chandy (1997) Curr. Opin. Biotechnol. 8:749-756).

Expression profiling

Microarrays are analytical tools used in bioanalysis. A microarray has a plurality of molecules spatially distributed over, and stably associated with, the surface of a solid support. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as gene sequencing, monitoring gene expression, gene mapping, bacterial identification, drug discovery, and combinatorial chemistry.

One area in particular in which microarrays find use is in gene expression analysis. Array technology can provide a simple way to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling

cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, condition, disease, or disorder.

Expression Profiling in Treatments for Cancer

Tumor cells stimulate the formation of stroma that secretes various mediators, such as growth factors, cytokines, and proteases, which are critical for tumor growth. For instance, serum tumor necrosis factor alpha (TNF-α) is increased in the circulation of patients with malignancy. Clinically, treatment with TNF-α, also called cachectin, in combination with Interferon-gamma (IFN-γ) may provide a successful approach to overcome the cellular heterogeneity of advanced breast tumors. TNF-α has been demonstrated to be antitumorigenic in MCF-7 cells by inducing apoptosis and inhibiting proliferation. TNF-α is produced by neutrophils, activated lymphocytes, macrophages, NK cells, LAK cells, astrocytes, endothelial cells, smooth muscle cells, and some transformed cells. TNF-α occurs as a secreted, soluble form and as a membrane-anchored form, both of which are biologically active. Two types of receptors for TNF-α have been described and virtually all cell types studied show the presence of one or both of these receptor types. TNF-α and TNF-β are extremely pleiotropic factors due to the ubiquity of their receptors, to their ability to activate multiple signal transduction pathways and to their ability to induce or suppress the expression of a wide number of genes. TNF-α and TNF-β play a critical role in mediation of the inflammatory response and in mediation of resistance to infections and tumor growth.

The cytokine interferon gamma (IFN-γ) induces growth arrest in normal human mammary epithelial cells by establishing a block during mid-G1 phase. IFN-γ inhibits the kinase activities of cdk2, cdk4 and cdk6 within 24 h of treatment. IFN-γ -mediated growth inhibition requires signal transducers and activators of transcription (STAT)-1 activation and may require induction of the cyclin-dependent kinase inhibitor p21. IFN-γ, possibly through the elevation of caspase-8 levels, sensitizes human breast tumor cells to a death receptor-mediated, mitochondria-operated pathway of apoptosis. IFN-γ, also known as Type II interferon or immune interferon, is produced primarily by T-lymphocytes and natural killer cells. IFN-γ exhibits antiproliferative, immunoregulatory and proinflammatory activities and is thus important in host defense mechanisms. IFN-γ induces the production of cytokines, and upregulates the expression of class I and II MHC antigens, Fc receptor, and leukocyte adhesion molecules. It modulates macrophage effector functions, influences isotype switching and potentiates the secretion of immunoglobulins by B cells. IFN-γ also augments TH1 cell expansion and may be required for TH1 cell differentiation. The IFN-γ receptor has been cloned and characterized, and is structurally related to the IL-10 receptor. It is present on almost all cell types except mature erythrocytes.

Breast cancer

Breast cancer is the most frequently diagnosed type of cancer in American women and the second most frequent cause of cancer death. The lifetime risk of an American woman developing breast cancer is 1 in 8, and one-third of women diagnosed with breast cancer die of the disease. A number of risk factors have been identified, including hormonal and genetic factors. One genetic defect associated with breast cancer results in a loss of heterozygosity (LOH) at multiple loci such as p53, Rb, BRCA1, and BRCA2. Another genetic defect is gene amplification involving genes such as c-myc and c-erbB2 (Her2-neu gene). Steroid and growth factor pathways are also altered in breast cancer, notably the estrogen, progesterone, and epidermal growth factor (EGF) pathways. Breast cancer evolves through a multi-step process whereby premalignant mammary epithelial cells undergo a relatively defined sequence of events leading to tumor formation. An early event in tumor development is ductal hyperplasia. Cells undergoing rapid neoplastic growth gradually progress to invasive carcinoma and become metastatic to the lung, bone, and potentially other organs. Variables that may influence the process of tumor progression and malignant transformation include genetic factors, environmental factors, growth factors, and hormones.

Lung cancer

Lung cancer is the leading cause of cancer death for men and the second leading cause of cancer death for women in the U.S. Lung cancers are divided into four histopathologically distinct groups. Three groups (squamous cell carcinoma, adenocarcinoma, and large cell carcinoma) are classified as non-small cell lung cancers (NSCLCs). The fourth group of cancers is referred to as small cell lung cancer (SCLC). Deletions on chromosome 3 are common in this disease and are thought to indicate the presence of a tumor suppressor gene in this region. Activating mutations in K-ras are commonly found in lung cancer and are the basis of one of the mouse models for the disease.

Colon cancer

While soft tissue sarcomas are relatively rare, more than 50% of new patients diagnosed with the disease will die from it. The molecular pathways leading to the development of sarcomas are relatively unknown, due to the rarity of the disease and variation in pathology. Colon cancer evolves through a multi-step process whereby pre-malignant colonocytes undergo a relatively defined sequence of events leading to tumor formation. Several factors participate in the process of tumor progression and malignant transformation including genetic factors, mutations, and selection.

To understand the nature of gene alterations in colorectal cancer, a number of studies have focused on the inherited syndromes. Familial adenomatous polyposis (FAP), is caused by mutations in the adenomatous polyposis coli gene (APC), resulting in truncated or inactive forms of the protein.

This tumor suppressor gene has been mapped to chromosome 5q. Hereditary nonpolyposis colorectal cancer (HNPCC) is caused by mutations in mis-match repair genes. Although hereditary colon cancer syndromes occur in a small percentage of the population and most colorectal cancers are considered sporadic, knowledge from studies of the hereditary syndromes can be generally applied. For instance, somatic mutations in APC occur in at least 80% of sporadic colon tumors. APC mutations are thought to be the initiating event in the disease. Other mutations occur subsequently. Approximately 50% of colorectal cancers contain activating mutations in ras, while 85% contain inactivating mutations in p53. Changes in all of these genes lead to gene expression changes in colon cancer.

Osteosarcoma

Osteosarcoma is the most common malignant bone tumor in children. Approximately 80% of patients present with non-metastatic disease. After the diagnosis is made by an initial biopsy, treatment involves the use of 3–4 courses of neoadjuvant chemotherapy before definitive surgery, followed by post-operative chemotherapy. With currently available treatment regimens, approximately 30–40% of patients with non-metastatic disease relapse after therapy. Currently, there is no prognostic factor that can be used at the time of initial diagnosis to predict which patients will have a high risk of relapse. The only significant prognostic factor predicting the outcome in a patient with non-metastatic osteosarcoma is the histopathologic response of the primary tumor resected at the time of definitive surgery. The degree of necrosis in the primary tumor is a reflection of the tumor response to neoadjuvant chemotherapy. A higher degree of necrosis (good or favorable response) is associated with a lower risk of relapse and a better outcome. Patients with a lower degree of necrosis (poor or unfavorable response) have a much higher risk of relapse and poor outcome even after complete resection of the primary tumor. Unfortunately, poor outcome cannot be altered despite modification of post-operative chemotherapy to account for the resistance of the primary tumor to neoadjuvant chemotherapy. Thus, there is an urgent need to identify prognostic factors that can be used at the time of diagnosis to recognize the subtypes of osteosarcomas that have various risks of relapse, so that more appropriate chemotherapy can be used at the outset to improve the outcome.

Ovarian cancer

Ovarian cancer is the leading cause of death from a gynecologic cancer. The majority of ovarian cancers are derived from epithelial cells, and 70% of patients with epithelial ovarian cancers present with late-stage disease. As a result, the long-term survival rates for this disease is very low. Identification of early-stage markers for ovarian cancer would significantly increase the survival rate.

Genetic variations involved in ovarian cancer development include mutation of p53 and microsatellite instability. Gene expression patterns likely vary when normal ovary is compared to ovarian tumors.

Immune response

Human peripheral blood mononuclear cells (PBMCs) can be classified into discrete cellular populations representing the major cellular components of the immune system. PBMCs contain about 52% lymphocytes (12% B lymphocytes, 40% T lymphocytes, 20% NK cells, monocytes, and 3% various cells that include dendritic cells and progenitor cells. The proportions, as well as the biology of these cellular components tend to vary slightly between healthy individuals, depending on factors such as age, gender, past medical history, and genetic background.

Tumor necrosis factor alpha (TNF-α), also called cachectin, is produced by neutrophils, activated lymphocytes, macrophages, NK cells, LAK cells, astrocytes, endothelial cells, smooth muscle cells, and some transformed cells. TNF-α occurs as a secreted, soluble form and a membrane-anchored form, both of which are biologically active. Two types of receptors for TNF-α have been described, and virtually all cell types studied show the presence of one or both of these receptor types. TNF- α and TNF-β are extremely pleiotropic factors due to the ubiquity of their receptors, their ability to activate multiple signal transduction pathways, and their ability to induce or suppress the expression of a wide number of genes. TNF-α and TNF-β play a critical role in mediation of the inflammatory response and in mediation of resistance to infections and tumor growth.

There is a need in the art for new compositions, including nucleic acids and proteins, for the diagnosis, prevention, and treatment of transport, neurological, muscle, immunological and cell proliferative disorders.

## SUMMARY OF THE INVENTION

Various embodiments of the invention provide purified polypeptides, transporters and ion channels, referred to collectively as 'TRICH' and individually as 'TRICH-1,' 'TRICH-2,' 'TRICH-3,' 'TRICH-4,' 'TRICH-5,' 'TRICH-6,' 'TRICH-7,' 'TRICH-8,' 'TRICH-9,' 'TRICH-10,' 'TRICH-11,' 'TRICH-12,' 'TRICH-13,' 'TRICH-14,' 'TRICH-15,' 'TRICH-16,' 'TRICH-17,' 'TRICH-18,' 'TRICH-19,' 'TRICH-20,' 'TRICH-21,' 'TRICH-22,' 'TRICH-23,' 'TRICH-24,' 'TRICH-25,' 'TRICH-26,' 'TRICH-27,' 'TRICH-28,' 'TRICH-29,' 'TRICH-30,' 'TRICH-31,' 'TRICH-32,' 'TRICH-33,' 'TRICH-34,' 'TRICH-35,' 'TRICH-36,' 'TRICH-37,' 'TRICH-38,' 'TRICH-39,' 'TRICH-40,' 'TRICH-41,' 'TRICH-42,' 'TRICH-43,' 'TRICH-44,' 'TRICH-45,' 'TRICH-46,' 'TRICH-47,' 'TRICH-48,' 'TRICH-49,' 'TRICH-50,' 'TRICH-51,' 'TRICH-52,' 'TRICH-53,'0'TRICH-54,' 'TRICH-55,' 'TRICH-56,' 'TRICH-57,' 'TRICH-58,' and 'TRICH-59 'and

methods for using these proteins and their encoding polynucleotides for the detection, diagnosis, and treatment of diseases and medical conditions. Embodiments also provide methods for utilizing the purified transporters and ion channels and/or their encoding polynucleotides for facilitating the drug discovery process, including determination of efficacy, dosage, toxicity, and pharmacology. Related embodiments provide methods for utilizing the purified transporters and ion channels and/or their encoding polynucleotides for investigating the pathogenesis of diseases and medical conditions.

An embodiment provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59. Another embodiment provides an isolated polypeptide comprising an amino acid sequence of SEQ ID NO:1-59.

Still another embodiment provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59. In another embodiment, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-59. In an alternative embodiment, the polynucleotide is selected from the group consisting of SEQ ID NO:60-118.

Still another embodiment provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

Another embodiment provides a cell transformed with the recombinant polynucleotide. Yet another embodiment provides a transgenic organism comprising the recombinant polynucleotide.

Another embodiment provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Yet another embodiment provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

Still yet another embodiment provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In other embodiments, the polynucleotide can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90%

identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex. In a related embodiment, the method can include detecting the amount of the hybridization complex. In still other embodiments, the probe can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Still yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof. In a related embodiment, the method can include detecting the amount of the amplified target polynucleotide or fragment thereof.

Another embodiment provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and a pharmaceutically acceptable excipient. In one embodiment, the composition can comprise an amino acid sequence selected from the group consisting of SEQ ID NO:1-59. Other embodiments provide a method of treating a disease or condition associated with decreased or abnormal expression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

Yet another embodiment provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. Another embodiment provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

Still yet another embodiment provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. Another embodiment provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with overexpression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

Another embodiment provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence

selected from the group consisting of SEQ ID NO:1-59. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

Yet another embodiment provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

Still yet another embodiment provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

Another embodiment provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID

NO:60-118, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide can comprise a fragment of a polynucleotide selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

## BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for full length polynucleotide and polypeptide embodiments of the invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog, and the PROTEOME database identification numbers and annotations of PROTEOME database homologs, for polypeptide embodiments of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide embodiments, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide embodiments, along with selected fragments of the polynucleotides.

Table 5 shows representative cDNA libraries for polynucleotide embodiments.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze polynucleotides and polypeptides, along with applicable descriptions, references, and threshold parameters.

Table 8 shows single nucleotide polymorphisms found in polynucleotide sequences of the invention, along with allele frequencies in different human populations.

## DESCRIPTION OF THE INVENTION

Before the present proteins, nucleic acids, and methods are described, it is understood that embodiments of the invention are not limited to the particular machines, instruments, materials, and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the invention.

As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with various embodiments of the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

**DEFINITIONS**

"TRICH" refers to the amino acid sequences of substantially purified TRICH obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of TRICH. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH participates.

An "allelic variant" is an alternative form of the gene encoding TRICH. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or

many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding TRICH include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as TRICH or a polypeptide with at least one functional characteristic of TRICH. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding TRICH, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide encoding TRICH. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent TRICH. Deliberate amino acid substitutions may be made on the basis of one or more similarities in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of TRICH is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" can refer to an oligopeptide, a peptide, a polypeptide, or a protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid. Amplification may be carried out using polymerase chain reaction (PCR) technologies or other nucleic acid amplification technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of TRICH. Antagonists may include proteins such as antibodies, anticalins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of

TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind TRICH polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-$NH_2$), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker (Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13).

The term "intramer" refers to an aptamer which is expressed *in vivo*. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl. Acad. Sci. USA 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a polynucleotide having a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic TRICH, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide" and a "composition comprising a given polypeptide" can refer to any composition containing the given polynucleotide or polypeptide. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotides encoding TRICH or fragments of TRICH may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (Accelrys, Burlington MA) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

| Original Residue | Conservative Substitution |
|---|---|
| Ala | Gly, Ser |
| Arg | His, Lys |
| Asn | Asp, Gln, His |
| Asp | Asn, Glu |
| Cys | Ala, Ser |
| Gln | Asn, Glu, His |
| Glu | Asp, Gln, His |
| Gly | Ala |
| His | Asn, Arg, Gln, Glu |
| Ile | Leu, Val |
| Leu | Ile, Val |
| Lys | Arg, Gln, Glu |
| Met | Leu, Ile |
| Phe | His, Met, Leu, Trp, Tyr |
| Ser | Cys, Thr |
| Thr | Ser, Val |
| Trp | Phe, Tyr |
| Tyr | His, Phe, Trp |
| Val | Ile, Leu, Thr |

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of TRICH or a polynucleotide encoding TRICH which can be identical in sequence to, but shorter in length than, the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from about 5 to about 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:60-118 can comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:60-118, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:60-118 can be employed in one or more embodiments of methods of the invention, for example, in hybridization and

amplification technologies and in analogous methods that distinguish SEQ ID NO:60-118 from related polynucleotides. The precise length of a fragment of SEQ ID NO:60-118 and the region of SEQ ID NO:60-118 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-59 is encoded by a fragment of SEQ ID NO:60-118. A fragment of SEQ ID NO:1-59 can comprise a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-59. For example, a fragment of SEQ ID NO:1-59 can be used as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-59. The precise length of a fragment of SEQ ID NO:1-59 and the region of SEQ ID NO:1-59 to which the fragment corresponds can be determined based on the intended purpose for the fragment using one or more analytical methods described herein or otherwise known in the art.

A "full length" polynucleotide is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, alternatively, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of identical nucleotide matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using one or more computer algorithms or programs known in the art or described herein. For example, percent identity can be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989; CABIOS 5:151-153) and in Higgins, D.G. et al. (1992; CABIOS 8:189-191). For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms which can be used is provided by the National Center for Biotechnology Information (NCBI) Basic

Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 11*

*Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of identical residue matches between at least two polypeptide sequences aligned using

a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide. The phrases "percent similarity" and "% similarity," as applied to polypeptide sequences, refer to the percentage of residue matches, including identical residue matches and conservative substitutions, between at least two polypeptide sequences aligned using a standardized algorithm. In contrast, conservative substitutions are not included in the calculation of percent identity between polypeptide sequences.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Open Gap: 11 and Extension Gap: 1 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 3*

*Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 $\mu$g/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point ($T_m$) for the specific sequence at a defined ionic strength and pH. The $T_m$ is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating $T_m$ and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press, Cold Spring Harbor NY, ch. 9).

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 $\mu$g/ml. Organic solvent, such as

formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acids by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., $C_0t$ or $R_0t$ analysis) or formed between one nucleic acid present in solution and another nucleic acid immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or polynucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of TRICH which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of TRICH which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, antibody, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of TRICH. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of TRICH.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or

synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an TRICH may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of TRICH.

"Probe" refers to nucleic acids encoding TRICH, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acids. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in, for example, Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring

Harbor Press, Cold Spring Harbor NY), Ausubel, F.M. et al. (1999; Short Protocols in Molecular Biology, 4th ed., John Wiley & Sons, New York NY), and Innis, M. et al. (1990; PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a nucleic acid that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook and Russell (*supra*). The term recombinant includes

nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA molecule, is composed of the same linear sequence of nucleotides as the reference DNA molecule with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing TRICH, nucleic acids encoding TRICH, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60%

free, preferably at least about 75% free, and most preferably at least about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. In another embodiment, the nucleic acid can be introduced by infection with a recombinant viral vector, such as a lentiviral vector (Lois, C. et al. (2002) Science 295:868-872). The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques

for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook and Russell (*supra*).

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotides that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity or sequence similarity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity or sequence similarity over a certain defined length of one of the polypeptides.

## THE INVENTION

Various embodiments of the invention include new human transporters and ion channels (TRICH), the polynucleotides encoding TRICH, and the use of these compositions for the diagnosis,

treatment, or prevention of transport, neurological, muscle, immunological and cell proliferative disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide embodiments of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown. Column 6 shows the Incyte ID numbers of physical, full length clones corresponding to the polypeptide and polynucleotide sequences of the invention. The full length clones encode polypeptides which have at least 95% sequence identity to the polypeptide sequences shown in column 3.

Table 2 shows sequences with homology to polypeptide embodiments of the invention as identified by BLAST analysis against the GenBank protein (genpept) database and the PROTEOME database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog and the PROTEOME database identification numbers (PROTEOME ID NO:) of the nearest PROTEOME database homologs. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank and PROTEOME database homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Accelrys, Burlington MA). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are transporters and ion channels. For example, SEQ ID NO:7 is 99% identical, from residue M1 to residue E300, to human acetylcholine receptor beta-subunit preprotein (GenBank ID g560155) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.9e-199, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:7 also has homology to the cholinergic receptor (nicotinic) beta 1 subunit, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:7 also contains a neurotransmitter-gated ion-channel ligand binding domain and a neurotransmitter-gated ion-channel transmembrane region as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein families/domains, and a cation transporter family protein domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based TIGRFAM database of conserved protein families/domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:7 is a cholinergic receptor subunit. In an alternative example, SEQ ID NO:41 is 97% identical, from residue M1 to residue T241, to human gamma-aminobutyric acidA receptor alpha 2 subunit (GenBank ID g386422) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.4e-215, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:41 also has homology to the alpha 2 subunit of the GABA-A receptor, a chloride channel that is the major inhibitory neurotransmitter receptor in the brain as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:41 also contains a neurotransmitter-gated ion channel ligand binding domain and a neurotransmitter-gated ion channel transmembrane domain, as well as a cation transporter family protein domain, as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM and TIGRFAM databases of conserved protein families/domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses, and BLAST analyses against the PRODOM and DOMO databases, provide further corroborative evidence that SEQ ID NO:41 is a GABA receptor. SEQ ID NO:1-6, SEQ ID NO:8-40, and SEQ ID NO:42-59 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-59 are described in Table 7.

As shown in Table 4, the full length polynucleotide embodiments were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two

types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5′) and stop (3′) positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide embodiments, and of fragments of the polynucleotides which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:60-118 or that distinguish between SEQ ID NO:60-118 and related polynucleotides.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotides. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL_*XXXXXX*_$N_1$_$N_2$_*YYYYY*_$N_3$_$N_4$ represents a "stitched" sequence in which *XXXXXX* is the identification number of the cluster of sequences to which the algorithm was applied, and *YYYYY* is the number of the prediction generated by the algorithm, and $N_{1,2,3...}$, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FL*XXXXXX*_g*AAAAA*_g*BBBBB*_1_*N* is a "stretched" sequence, with *XXXXXX* being the Incyte project identification number, g*AAAAA* being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, g*BBBBB* being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used

as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, g*BBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

| **Prefix** | **Type of analysis and/or examples of programs** |
|---|---|
| GNN, GFG, ENST | Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK). |
| GBI | Hand-edited analysis of genomic sequences. |
| FL | Stitched or stretched genomic sequences (see Example V). |
| INCY | Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript. |

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotides which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotides. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

Table 8 shows single nucleotide polymorphisms (SNPs) found in polynucleotide sequences of the invention, along with allele frequencies in different human populations. Columns 1 and 2 show the polynucleotide sequence identification number (SEQ ID NO:) and the corresponding Incyte project identification number (PID) for polynucleotides of the invention. Column 3 shows the Incyte identification number for the EST in which the SNP was detected (EST ID), and column 4 shows the identification number for the SNP (SNP ID). Column 5 shows the position within the EST sequence at which the SNP is located (EST SNP), and column 6 shows the position of the SNP within the full-length polynucleotide sequence (CB1 SNP). Column 7 shows the allele found in the EST sequence.

Columns 8 and 9 show the two alleles found at the SNP site. Column 10 shows the amino acid encoded by the codon including the SNP site, based upon the allele found in the EST. Columns 11-14 show the frequency of allele 1 in four different human populations. An entry of n/d (not detected) indicates that the frequency of allele 1 in the population was too low to be detected, while n/a (not available) indicates that the allele frequency was not determined for the population.

The invention also encompasses TRICH variants. Various embodiments of TRICH variants can have at least about 80%, at least about 90%, or at least about 95% amino acid sequence identity to the TRICH amino acid sequence, and can contain at least one functional or structural characteristic of TRICH.

Various embodiments also encompass polynucleotides which encode TRICH. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:60-118, which encodes TRICH. The polynucleotide sequences of SEQ ID NO:60-118, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses variants of a polynucleotide encoding TRICH. In particular, such a variant polynucleotide will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a polynucleotide encoding TRICH. A particular aspect of the invention encompasses a variant of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NO:60-118 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:60-118. Any one of the polynucleotide variants described above can encode a polypeptide which contains at least one functional or structural characteristic of TRICH.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide encoding TRICH. A splice variant may have portions which have significant sequence identity to a polynucleotide encoding TRICH, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to a polynucleotide encoding TRICH over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100%

polynucleotide sequence identity to portions of the polynucleotide encoding TRICH. For example, a polynucleotide comprising a sequence of SEQ ID NO:63 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:66; and a polynucleotide comprising a sequence of SEQ ID NO:64 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:68. In an alternative example, a polynucleotide comprising a sequence of SEQ ID NO:97, a polynucleotide comprising a sequence of SEQ ID NO:98, a polynucleotide comprising a sequence of SEQ ID NO:99, a polynucleotide comprising a sequence of SEQ ID NO:100, a polynucleotide comprising a sequence of SEQ ID NO:101, a polynucleotide comprising a sequence of SEQ ID NO:102, and a polynucleotide comprising a sequence of SEQ ID NO:114 are all splice variants of each other. In a further example, a polynucleotide comprising a sequence of SEQ ID NO:93 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:94, a polynucleotide comprising a sequence of SEQ ID NO:106 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:107, and a polynucleotide comprising a sequence of SEQ ID NO:116 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:117. In addition, a polynucleotide comprising a sequence of SEQ ID NO:60 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:79, a polynucleotide comprising a sequence of SEQ ID NO:67 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:84, a polynucleotide comprising a sequence of SEQ ID NO:71 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:75, and a polynucleotide comprising a sequence of SEQ ID NO:73 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:81. Any one of the splice variants described above can encode a polypeptide which contains at least one functional or structural characteristic of TRICH.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding TRICH, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring TRICH, and all such variations are to be considered as being specifically disclosed.

Although polynucleotides which encode TRICH and its variants are generally capable of hybridizing to polynucleotides encoding naturally occurring TRICH under appropriately selected conditions of stringency, it may be advantageous to produce polynucleotides encoding TRICH or its

derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding TRICH and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of polynucleotides which encode TRICH and TRICH derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic polynucleotide may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a polynucleotide encoding TRICH or any fragment thereof.

Embodiments of the invention can also include polynucleotides that are capable of hybridizing to the claimed polynucleotides, and, in particular, to those having the sequences shown in SEQ ID NO:60-118 and fragments thereof, under various conditions of stringency (Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511). Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Biosciences, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Invitrogen, Carlsbad CA). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Amersham Biosciences), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art (Ausubel et al., *supra*, ch. 7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853).

The nucleic acids encoding TRICH may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as

promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art (Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotides or fragments thereof which encode TRICH may be cloned in recombinant DNA molecules that direct expression of TRICH, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other polynucleotides which encode substantially the same or a functionally equivalent polypeptides may be produced and used to express TRICH.

The polynucleotides of the invention can be engineered using methods generally known in the art in order to alter TRICH-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of TRICH, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, polynucleotides encoding TRICH may be synthesized, in whole or in part, using one or more chemical methods well known in the art (Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232). Alternatively, TRICH itself or a fragment thereof may be synthesized using chemical methods known in the art. For example, peptide synthesis can be performed using various solution-phase or

solid-phase techniques (Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J.Y. et al. (1995) Science 269:202-204). Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of TRICH, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography (Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421). The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing (Creighton, *supra*, pp. 28-53).

In order to express a biologically active TRICH, the polynucleotides encoding TRICH or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotides encoding TRICH. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of polynucleotides encoding TRICH. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where a polynucleotide sequence encoding TRICH and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

Methods which are well known to those skilled in the art may be used to construct expression vectors containing polynucleotides encoding TRICH and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination (Sambrook and Russell, *supra*, ch. 1-4, and 8; Ausubel et al., *supra*, ch. 1, 3, and 15).

A variety of expression vector/host systems may be utilized to contain and express polynucleotides encoding TRICH. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems (Sambrook and Russell, *supra*; Ausubel et al., *supra*; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355). Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of polynucleotides to the targeted organ, tissue, or cell population (Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5:350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Buller, R.M. et al. (1985) Nature 317:813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31:219-226; Verma, I.M. and N. Somia (1997) Nature 389:239-242). The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotides encoding TRICH. For example, routine cloning, subcloning, and propagation of polynucleotides encoding TRICH can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Invitrogen). Ligation of polynucleotides encoding TRICH into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509). When large quantities of TRICH are needed, e.g. for the production of antibodies, vectors which direct high level expression of TRICH may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of TRICH. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH

promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign polynucleotide sequences into the host genome for stable propagation (Ausubel et al., *supra*; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184).

Plant systems may also be used for expression of TRICH. Transcription of polynucleotides encoding TRICH may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection (The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196).

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, polynucleotides encoding TRICH may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses TRICH in host cells (Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355).

For long term production of recombinant proteins in mammalian systems, stable expression of TRICH in cell lines is preferred. For example, polynucleotides encoding TRICH can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer

resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively (Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823). Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14). Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051). Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β-glucuronidase and its substrate β-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding TRICH is inserted within a marker gene sequence, transformed cells containing polynucleotides encoding TRICH can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding TRICH under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the polynucleotide encoding TRICH and that express TRICH may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of TRICH using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques

include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on TRICH is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art (Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding TRICH include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, polynucleotides encoding TRICH, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Biosciences, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with polynucleotides encoding TRICH may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode TRICH may be designed to contain signal sequences which direct secretion of TRICH through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted polynucleotides or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for

post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant polynucleotides encoding TRICH may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric TRICH protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of TRICH activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the TRICH encoding sequence and the heterologous protein sequence, so that TRICH may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In another embodiment, synthesis of radiolabeled TRICH may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, $^{35}$S-methionine.

TRICH, fragments of TRICH, or variants of TRICH may be used to screen for compounds that specifically bind to TRICH. One or more test compounds may be screened for specific binding to TRICH. In various embodiments, 1, 2, 3, 4, 5, 10, 20, 50, 100, or 200 test compounds can be screened for specific binding to TRICH. Examples of test compounds can include antibodies, anticalins, oligonucleotides, proteins (e.g., ligands or receptors), or small molecules.

In related embodiments, variants of TRICH can be used to screen for binding of test compounds, such as antibodies, to TRICH, a variant of TRICH, or a combination of TRICH and/or one or more variants TRICH. In an embodiment, a variant of TRICH can be used to screen for compounds that bind to a variant of TRICH, but not to TRICH having the exact sequence of a sequence of SEQ ID NO:1-59. TRICH variants used to perform such screening can have a range of about 50% to about 99% sequence identity to TRICH, with various embodiments having 60%, 70%, 75%, 80%, 85%, 90%, and 95% sequence identity.

In an embodiment, a compound identified in a screen for specific binding to TRICH can be closely related to the natural ligand of TRICH, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner (Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2):Chapter 5). In another embodiment, the compound thus identified can be a natural ligand of a receptor TRICH (Howard, A.D. et al. (2001) Trends Pharmacol. Sci.22:132-140; Wise, A. et al. (2002) Drug Discovery Today 7:235-246).

In other embodiments, a compound identified in a screen for specific binding to TRICH can be closely related to the natural receptor to which TRICH binds, at least a fragment of the receptor, or a fragment of the receptor including all or a portion of the ligand binding site or binding pocket. For example, the compound may be a receptor for TRICH which is capable of propagating a signal, or a decoy receptor for TRICH which is not capable of propagating a signal (Ashkenazi, A. and V.M. Divit (1999) Curr. Opin. Cell Biol. 11:255-260; Mantovani, A. et al. (2001) Trends Immunol. 22:328-336). The compound can be rationally designed using known techniques. Examples of such techniques include those used to construct the compound etanercept (ENBREL; Amgen Inc., Thousand Oaks CA), which is efficacious for treating rheumatoid arthritis in humans. Etanercept is an engineered p75 tumor necrosis factor (TNF) receptor dimer linked to the Fc portion of human $IgG_1$ (Taylor, P.C. et al. (2001) Curr. Opin. Immunol. 13:611-616).

In one embodiment, two or more antibodies having similar or, alternatively, different specificities can be screened for specific binding to TRICH, fragments of TRICH, or variants of TRICH. The binding specificity of the antibodies thus screened can thereby be selected to identify particular fragments or variants of TRICH. In one embodiment, an antibody can be selected such that its binding specificity allows for preferential identification of specific fragments or variants of TRICH. In another embodiment, an antibody can be selected such that its binding specificity allows for preferential diagnosis of a specific disease or condition having increased, decreased, or otherwise abnormal production of TRICH.

In an embodiment, anticalins can be screened for specific binding to TRICH, fragments of TRICH, or variants of TRICH. Anticalins are ligand-binding proteins that have been constructed based on a lipocalin scaffold (Weiss, G.A. and H.B. Lowman (2000) Chem. Biol. 7:R177-R184; Skerra, A. (2001) J. Biotechnol. 74:257-275). The protein architecture of lipocalins can include a beta-barrel having eight antiparallel beta-strands, which supports four loops at its open end. These loops form the natural ligand-binding site of the lipocalins, a site which can be re-engineered *in vitro* by amino acid substitutions to impart novel binding specificities. The amino acid substitutions can be made using methods known in the art or described herein, and can include conservative substitutions (e.g., substitutions that do not alter binding specificity) or substitutions that modestly, moderately, or significantly alter binding specificity.

In one embodiment, screening for compounds which specifically bind to, stimulate, or inhibit TRICH involves producing appropriate cells which express TRICH, either as a secreted protein or on the cell membrane. Preferred cells can include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing TRICH or cell membrane fractions which contain TRICH are then contacted with a test compound and binding, stimulation, or inhibition of activity of either TRICH or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with TRICH, either in solution or affixed to a solid support, and detecting the binding of TRICH to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

An assay can be used to assess the ability of a compound to bind to its natural ligand and/or to inhibit the binding of its natural ligand to its natural receptors. Examples of such assays include radio-labeling assays such as those described in U.S. Patent No. 5,914,236 and U.S. Patent No. 6,372,724. In a related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a receptor) to improve or alter its ability to bind to its natural ligands (Matthews, D.J. and J.A. Wells. (1994) Chem. Biol. 1:25-30). In another related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a ligand) to improve or alter its ability to bind to its natural receptors (Cunningham, B.C. and J.A. Wells (1991) Proc. Natl. Acad. Sci. USA 88:3407-3411; Lowman, H.B. et al. (1991) J. Biol. Chem. 266:10982-10988).

TRICH, fragments of TRICH, or variants of TRICH may be used to screen for compounds that modulate the activity of TRICH. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for TRICH activity, wherein TRICH is combined with at least one test compound, and the activity of TRICH in the presence of a test compound is compared with the activity of TRICH in the absence of the test compound. A change in the activity of TRICH in the presence of the test compound is indicative of a compound that modulates the activity of TRICH. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising TRICH under conditions suitable for TRICH activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of TRICH may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding TRICH or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease (see, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337). For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding TRICH may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding TRICH can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding TRICH is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress TRICH, e.g., by secreting TRICH in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

**THERAPEUTICS**

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of TRICH and transporters and ion channels. In addition, examples of tissues expressing TRICH can be found in Table 6 and can also be found in Example XI. Therefore, TRICH appears to play a role in transport, neurological, muscle, immunological and cell proliferative disorders. In the treatment of disorders associated with increased TRICH expression or activity, it is desirable to decrease the expression or activity of TRICH. In the treatment of disorders associated with decreased TRICH expression or activity, it is desirable to increase the expression or activity of TRICH.

Therefore, in one embodiment, TRICH or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy,

sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, glycogen storage disease, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, pseudohypoaldosteronism type 1, Liddle's syndrome, cystinuria, iminoglycinuria, Hartup disease, Fanconi disease, and Bartter syndrome; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, hemiplegic migraine, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, acid maltase deficiency (AMD, also known as Pompe's disease), generalized myotonia, and myotonia congenita; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease,

adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing TRICH or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified TRICH in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of TRICH may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those listed above.

In a further embodiment, an antagonist of TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH. Examples of such disorders include, but are not limited to, those transport, neurological, muscle, immunological and cell

proliferative disorders described above. In one aspect, an antibody which specifically binds TRICH may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express TRICH.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH including, but not limited to, those described above.

In other embodiments, any protein, agonist, antagonist, antibody, complementary sequence, or vector embodiments may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of TRICH may be produced using methods which are generally known in the art. In particular, purified TRICH may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind TRICH. Antibodies to TRICH may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. In an embodiment, neutralizing antibodies (i.e., those which inhibit dimer formation) can be used therapeutically. Single chain antibodies (e.g., from camels or llamas) may be potent enzyme inhibitors and may have application in the design of peptide mimetics, and in the development of immuno-adsorbents and biosensors (Muyldermans, S. (2001) J. Biotechnol. 74:277-302).

For the production of antibodies, various hosts including goats, rabbits, rats, mice, camels, dromedaries, llamas, humans, and others may be immunized by injection with TRICH or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to TRICH have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are substantially identical to a portion of the amino acid sequence of the natural protein. Short stretches of TRICH amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to TRICH may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce TRICH-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for TRICH may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 246:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between TRICH and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering TRICH epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for TRICH. Affinity is expressed as an association constant, $K_a$, which is defined as the molar concentration of TRICH-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The $K_a$ determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple TRICH epitopes, represents the average affinity, or avidity, of the antibodies for TRICH. The $K_a$ determined for a preparation of monoclonal antibodies, which are monospecific for a particular TRICH epitope, represents a true measure of affinity. High-affinity antibody preparations with $K_a$ ranging from about $10^9$ to $10^{12}$ L/mole are preferred for use in immunoassays in which the TRICH-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with $K_a$ ranging from about $10^6$ to $10^7$ L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of TRICH, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of TRICH-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available (Catty, *supra*; Coligan et al., *supra*).

In another embodiment of the invention, polynucleotides encoding TRICH, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA,

RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding TRICH. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding TRICH (Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press, Totawa NJ).

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein (Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102:469-475; Scanlon, K.J. et al. (1995) 9:1288-1296). Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors (Miller, A.D. (1990) Blood 76:271; Ausubel et al., *supra*; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63:323-347). Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art (Rossi, J.J. (1995) Br. Med. Bull. 51:217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87:1308-1315; Morris, M.C. et al. (1997) Nucleic Acids Res. 25:2730-2736).

In another embodiment of the invention, polynucleotides encoding TRICH may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in TRICH expression or regulation causes disease, the expression

of TRICH from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in TRICH are treated by constructing mammalian expression vectors encoding TRICH and introducing these vectors by mechanical means into TRICH-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J.-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of TRICH include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). TRICH may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding TRICH from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to TRICH expression are treated by constructing a retrovirus vector consisting of (i) the

polynucleotide encoding TRICH under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In an embodiment, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding TRICH to cells which have one or more genetic abnormalities with respect to the expression of TRICH. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999; Annu. Rev. Nutr. 19:511-544) and Verma, I.M. and N. Somia (1997; Nature 18:389:239-242).

In another embodiment, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TRICH to target cells which have one or more genetic abnormalities with respect to the expression of TRICH. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TRICH to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with

ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999; J. Virol. 73:519-532) and Xu, H. et al. (1994; Dev. Biol. 163:152-161). The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another embodiment, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding TRICH to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for TRICH into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TRICH-coding RNAs and the synthesis of high levels of TRICH in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of TRICH into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of RNA molecules encoding TRICH.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA molecules encoding TRICH. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be

extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

In other embodiments of the invention, the expression of one or more selected polynucleotides of the present invention can be altered, inhibited, decreased, or silenced using RNA interference (RNAi) or post-transcriptional gene silencing (PTGS) methods known in the art. RNAi is a post-transcriptional mode of gene silencing in which double-stranded RNA (dsRNA) introduced into a targeted cell specifically suppresses the expression of the homologous gene (i.e., the gene bearing the sequence complementary to the dsRNA). This effectively knocks out or substantially reduces the expression of the targeted gene. PTGS can also be accomplished by use of DNA or DNA fragments as well. RNAi methods are described by Fire, A. et al. (1998; Nature 391:806-811) and Gura, T. (2000; Nature 404:804-808). PTGS can also be initiated by introduction of a complementary segment of DNA into the selected tissue using gene delivery and/or viral vector delivery methods described herein or known in the art.

RNAi can be induced in mammalian cells by the use of small interfering RNA also known as siRNA. SiRNA are shorter segments of dsRNA (typically about 21 to 23 nucleotides in length) that result *in vivo* from cleavage of introduced dsRNA by the action of an endogenous ribonuclease. SiRNA appear to be the mediators of the RNAi effect in mammals. The most effective siRNAs appear to be 21 nucleotide dsRNAs with 2 nucleotide 3' overhangs. The use of siRNA for inducing RNAi in mammalian cells is described by Elbashir, S.M. et al. (2001; Nature 411:494-498).

SiRNA can either be generated indirectly by introduction of dsRNA into the targeted cell, or directly by mammalian transfection methods and agents described herein or known in the art (such as liposome-mediated transfection, viral vector methods, or other polynucleotide delivery/introductory methods). Suitable SiRNAs can be selected by examining a transcript of the target polynucleotide (e.g., mRNA) for nucleotide sequences downstream from the AUG start codon and recording the occurrence of each nucleotide and the 3' adjacent 19 to 23 nucleotides as potential siRNA target sites, with sequences having a 21 nucleotide length being preferred. Regions to be avoided for target siRNA sites include the 5' and 3' untranslated regions (UTRs) and regions near the start codon (within 75 bases), as these may be richer in regulatory protein binding sites. UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP endonuclease complex. The selected target sites for siRNA can then be compared to the appropriate genome database (e.g., human, etc.) using BLAST or other sequence comparison algorithms known in the art. Target

sequences with significant homology to other coding sequences can be eliminated from consideration. The selected SiRNAs can be produced by chemical synthesis methods known in the art or by *in vitro* transcription using commercially available methods and kits such as the SILENCER siRNA construction kit (Ambion, Austin TX).

In alternative embodiments, long-term gene silencing and/or RNAi effects can be induced in selected tissue using expression vectors that continuously express siRNA. This can be accomplished using expression vectors that are engineered to express hairpin RNAs (shRNAs) using methods known in the art (see, e.g., Brummelkamp, T.R. et al. (2002) Science 296:550-553; and Paddison, P.J. et al. (2002) Genes Dev. 16:948-958). In these and related embodiments, shRNAs can be delivered to target cells using expression vectors known in the art. An example of a suitable expression vector for delivery of siRNA is the PSILENCER1.0-U6 (circular) plasmid (Ambion). Once delivered to the target tissue, shRNAs are processed *in vivo* into siRNA-like molecules capable of carrying out gene-specific silencing.

In various embodiments, the expression levels of genes targeted by RNAi or PTGS methods can be determined by assays for mRNA and/or protein analysis. Expression levels of the mRNA of a targeted gene, can be determined by northern analysis methods using, for example, the NORTHERNMAX-GLY kit (Ambion); by microarray methods; by PCR methods; by real time PCR methods; and by other RNA/polynucleotide assays known in the art or described herein. Expression levels of the protein encoded by the targeted gene can be determined by Western analysis using standard techniques known in the art.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding TRICH. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression.. Thus, in the treatment of disorders associated with increased TRICH expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding TRICH may be therapeutically useful, and in the treatment of disorders associated with decreased TRICH expression or activity, a compound which specifically promotes expression of the polynucleotide encoding TRICH may be therapeutically useful.

In various embodiments, one or more test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding TRICH is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding TRICH are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding TRICH. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art (Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466).

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of TRICH, antibodies to TRICH, and mimetics, agonists, antagonists, or inhibitors of TRICH.

In various embodiments, the compositions described herein, such as pharmaceutical compositions, may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery allows administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising TRICH or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, TRICH or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example TRICH or fragments thereof, antibodies of TRICH, and agonists, antagonists or inhibitors of TRICH, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the $ED_{50}$ (the dose therapeutically effective in 50% of the population) or $LD_{50}$ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the $LD_{50}/ED_{50}$ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the $ED_{50}$ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 $\mu$g to 100,000 $\mu$g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

# DIAGNOSTICS

In another embodiment, antibodies which specifically bind TRICH may be used for the diagnosis of disorders characterized by expression of TRICH, or in assays to monitor patients being treated with TRICH or agonists, antagonists, or inhibitors of TRICH. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TRICH include methods which utilize the antibody and a label to detect TRICH in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring TRICH, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of TRICH expression. Normal or standard values for TRICH expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to TRICH under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of TRICH expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, polynucleotides encoding TRICH may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotides, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of TRICH may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of TRICH, and to monitor regulation of TRICH levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotides, including genomic sequences, encoding TRICH or closely related molecules may be used to identify nucleic acid sequences which encode TRICH. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding TRICH, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the TRICH encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:60-118 or from genomic sequences including promoters, enhancers, and introns of the TRICH gene.

Means for producing specific hybridization probes for polynucleotides encoding TRICH include the cloning of polynucleotides encoding TRICH or TRICH derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as $^{32}P$ or $^{35}S$, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotides encoding TRICH may be used for the diagnosis of disorders associated with expression of TRICH. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, glycogen storage disease, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, pseudohypoaldosteronism type 1, Liddle's syndrome, cystinuria, iminoglycinuria, Hartup disease, Fanconi disease, and Bartter syndrome; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial

thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, hemiplegic migraine, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, acid maltase deficiency (AMD, also known as Pompe's disease), generalized myotonia, and myotonia congenita; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial,

fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. Polynucleotides encoding TRICH may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered TRICH expression. Such qualitative or quantitative methods are well known in the art.

In a particular embodiment, polynucleotides encoding TRICH may be used in assays that detect the presence of associated disorders, particularly those mentioned above. Polynucleotides complementary to sequences encoding TRICH may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of polynucleotides encoding TRICH in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of TRICH, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding TRICH, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the

patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding TRICH may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding TRICH, or a fragment of a polynucleotide complementary to the polynucleotide encoding TRICH, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from polynucleotides encoding TRICH may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from polynucleotides encoding TRICH are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass

spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

SNPs may be used to study the genetic basis of human disease. For example, at least 16 common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase pathway. Analysis of the distribution of SNPs in different populations is useful for investigating genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations and their migrations (Taylor, J.G. et al. (2001) Trends Mol. Med. 7:507-512; Kwok, P.-Y. and Z. Gu (1999) Mol. Med. Today 5:538-543; Nowotny, P. et al. (2001) Curr. Opin. Neurobiol. 11:637-641).

Methods which may also be used to quantify the expression of TRICH include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves (Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment

regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, TRICH, fragments of TRICH, or antibodies specific for TRICH may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time (Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484; hereby expressly incorporated by reference herein). Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*, as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different

compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity (see, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm). Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In an embodiment, the toxicity of a test compound can be assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another embodiment relates to the use of the polypeptides disclosed herein to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous

amino acid residues, to the polypeptide sequences of interest. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for TRICH to quantify the levels of TRICH expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art (Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/25116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and thoroughly described in Schena, M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

In another embodiment of the invention, nucleic acid sequences encoding TRICH may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; Trask, B.J. (1991) Trends Genet. 7:149-154). Once mapped, the nucleic acid sequences may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP) (Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357).

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data (Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968). Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding TRICH on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

*In situ* hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery

techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation (Gatti, R.A. et al. (1988) Nature 336:577-580). The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, TRICH, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between TRICH and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest (Geysen, et al. (1984) PCT application WO84/03564). In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with TRICH, or fragments thereof, and washed. Bound TRICH is then detected by methods well known in the art. Purified TRICH can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding TRICH specifically compete with a test compound for binding TRICH. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TRICH.

In additional embodiments, the nucleotide sequences which encode TRICH may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/368,840, and U.S. Ser. No. 60/375,637, are hereby expressly incorporated by reference.

## EXAMPLES

### I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Invitrogen), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Invitrogen), using the recommended procedures or similar methods known in the art (Ausubel et al., *supra*, ch. 5). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Biosciences) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Invitrogen, Carlsbad CA), PCDNA2.1 plasmid (Invitrogen), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics),

or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Invitrogen.

**II. Isolation of cDNA Clones**

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

**III. Sequencing and Analysis**

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Biosciences or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Amersham Biosciences); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (Ausubel et al., *supra*, ch. 7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Candida albicans* (Incyte Genomics, Palo Alto CA); hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM (Haft, D.H. et al. (2001) Nucleic Acids Res. 29:41-43); and HMM-based protein domain databases such as SMART (Schultz, J. et al. (1998) Proc. Natl. Acad. Sci. USA 95:5857-5864; Letunic, I. et al. (2002) Nucleic Acids Res. 30:242-244). (HMM is a probabilistic approach which analyzes consensus primary structures of gene families; see, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM; and HMM-based protein domain databases such as SMART. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (MiraiBio, Alameda CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:60-118. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

**IV. Identification and Editing of Coding Sequences from Genomic DNA**

Putative transporters and ion channels were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94; Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode transporters and ion channels, the encoded polypeptides were analyzed by querying against PFAM models for transporters and ion channels. Potential transporters and ion channels were also identified by homology to Incyte cDNA sequences that had been annotated as transporters and ion channels. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA

sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

## V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions

may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

**VI. Chromosomal Mapping of TRICH Encoding Polynucleotides**

The sequences which were used to assemble SEQ ID NO:60-118 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:60-118 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

**VII. Analysis of Polynucleotide Expression**

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook and Russell, *supra*, ch. 7; Ausubel et al., *supra*, ch. 4).

Analogous computer techniques applying BLAST were used to search for identical or related molecules in databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much

faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{5 \text{ x minimum } \{\text{length(Seq. 1), length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotides encoding TRICH are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific

expression of cDNA encoding TRICH. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

**VIII. Extension of TRICH Encoding Polynucleotides**

Full length polynucleotides are produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing $Mg^{2+}$, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Biosciences), ELONGASE enzyme (Invitrogen), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 $\mu$l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 $\mu$l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 $\mu$l to 10 $\mu$l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Biosciences). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Biosciences), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Biosciences) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Biosciences) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotides are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

## IX. Identification of Single Nucleotide Polymorphisms in TRICH Encoding Polynucleotides

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:60-118 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants. An automated procedure of advanced chromosome analysis analysed the original chromatogram files

in the vicinity of the putative SNP. Clone error filters used statistically generated algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

Certain SNPs were selected for further characterization by mass spectrometry using the high throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezualan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed no allelic variance in this population were not further tested in the other three populations.

**X. Labeling and Use of Individual Hybridization Probes**

Hybridization probes derived from SEQ ID NO:60-118 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 $\mu$Ci of [$\gamma$-$^{32}$P] adenosine triphosphate (Amersham Biosciences), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Biosciences). An aliquot containing $10^7$ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate.

Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

## XI. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing; see, e.g., Baldeschweiler et al., *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena, M., ed. (1999) DNA Microarrays: A Practical Approach, Oxford University Press, London). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements (Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31).

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

### Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/$\mu$l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/$\mu$l RNase inhibitor, 500 $\mu$M dATP, 500 $\mu$M dGTP, 500 $\mu$M dTTP, 40 $\mu$M dCTP, 40 $\mu$M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Biosciences). The reverse transcription

reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte Genomics). Specific control poly(A)$^+$ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37° C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85° C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech, Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 $\mu$l 5X SSC/0.2% SDS.

**Microarray Preparation**

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 $\mu$g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Biosciences).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma-Aldrich, St. Louis MO) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1 $\mu$l of the array element DNA, at an average concentration of 100 ng/$\mu$l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60° C followed by washes in 0.2% SDS and distilled water as before.

**Hybridization**

Hybridization reactions contain 9 $\mu$l of sample mixture consisting of 0.2 $\mu$g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm$^2$ coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 $\mu$l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45° C in a second wash buffer (0.1X SSC), and dried.

**Detection**

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte Genomics). Array elements that exhibit at least about a two-fold change in expression, a signal-to-background ratio of at least about 2.5, and an element spot size of at least about 40%, are considered to be differentially expressed.

**Expression**

Breast cancer

For example, SEQ ID NO:85 showed decreased expression in nonmalignant breast adenocarcinoma cells treated with serum tumor necrosis factor alpha (TNF-α) versus untreated nonmalignant breast adenocarcinoma cells as determined by microarray analysis. MCF7 is a nonmalignant breast adenocarcinoma cell line isolated from the pleural effusion of a 69- year-old female. MCF7 has retained characteristics of the mammary epithelium such as the ability to process estradiol via cytoplasmic estrogen receptors and the capacity to form domes in culture. MCF7 cells were treated with TNF-α for 1, 4, 8, 12, 24, 36, 48, and 72 hours. Treated cells were compared to untreated cells kept in culture for the same amount of time. The expression of SEQ ID NO:85 was reduced by at least two-fold at later time points. In addition, SEQ ID NO:85 showed decreased expression in breast carcinoma cells treated with interferon gamma (IFNγ) versus untreated breast carcinoma cells. T-47D is a breast carcinoma cell line isolated from a pleural effusion obtained from a 54-year-old female with an infiltrating ductal carcinoma of the breast. T-47D cells were treated with 200 ng/ml IFNγ for 1, 4, 8, 24, 48 hours and 3 days. These treated cells were compared to untreated cells. The expression of SEQ ID NO:85 was reduced by at least two-fold at later time points.

In a further example, SEQ ID NO:88 showed differential expression in breast cell carcinoma cells versus nonmalignant mammary epithelial cells as determined by microarray analysis. Gene expression profiles of nonmalignant mammary epithelial cells were compared to the gene expression profile of a breast carcinoma line. The cells were grown in defined serum-free H14 medium to 70-80% confluence prior to RNA harvest. Cell lines compared include T-47D, a breast carcinoma cell line isolated from a pleural effusion obtained from a 54-year-old female with an infiltrating ductal carcinoma of the breast versus MCF-10A, a breast mammary gland cell line isolated from a 36-year-old woman with fibrocystic breast disease, and HMEC, a primary breast epithelial cell line isolated from a normal donor. The expression of SEQ ID NO:88 was increased by at least two-fold in T-47D cells as compared to either HMEC or MCF-10A cells.

In a further example, SEQ ID NO:112 showed differential expression in breast tumor tissue as compared to normal breast tissue from the same donor as determined by microarray analysis. Tumor from the right breast was compared to grossly uninvolved breast tissue from the same donor, a 43-year-old female diagnosed with invasive lobular carcinoma *in situ*. The expression of SEQ ID NO:112 was decreased by at least two-fold in the tumor tissue as compared to the matched non-tumor tissue.

In a further example, SEQ ID NO:113 showed differential expression in breast cancer cell lines as compared to non-cancerous breast epithelial cell lines as determined by microarray analysis.. Cell lines compared included: a) BT-20, a breast carcinoma cell line derived *in vitro* from the cells emigrating out of thin slices of tumor mass isolated from a 74-year-old female, b) BT-474, a breast ductal carcinoma cell line that was isolated from a solid, invasive ductal carcinoma of the breast obtained from a 60-year-old woman, c) BT-483, a breast ductal carcinoma cell line that was isolated from a papillary invasive ductal tumor obtained from a 23-year-old normal, menstruating, parous female with a family history of breast cancer, d) Hs 578T, a breast ductal carcinoma cell line isolated from a 74-year-old female with breast carcinoma, e) MCF7, a nonmalignant breast adenocarcinoma cell line isolated from the pleural effusion of a 69-year-old female, f) MCF-10A, a breast mammary gland (luminal ductal characteristics) cell line isolated from a 36-year-old woman with fibrocystic breast disease, g) MDA-MB-468, a breast adenocarcinoma cell line isolated from the pleural effusion of a 51-year-old female with metastatic adenocarcinoma of the breast, and h) HMEC, a primary breast epithelial cell line isolated from a normal donor. Expression of SEQ ID NO: 113 was decreased by at least two-fold in the BT-474 and BT-483 breast cancer cell lines as compared to the

non-malignant HMEC cells. Therefore, SEQ ID NO: 113 is useful in monitoring treatment of, and diagnostic assays for, breast cancer.

Therefore, in various embodiments, SEQ ID NO:85, SEQ ID NO:88, and SEQ ID NO: 112-113 can each be used for one or more of the following: i) monitoring treatment of breast adenocarcinoma and other cell proliferative disorders, ii) diagnostic assays for breast adenocarcinoma and other cell proliferative disorders, and iii) developing therapeutics and/or other treatments for breast adenocarcinoma and other cell proliferative disorders.

<u>Lung cancer</u>

In another example, SEQ ID NO:85 showed increased expression in lung tumor tissue versus normal lung tissue. Normal lung tissue from a 68 year-old female was compared to lung tumor from the same donor (Roy Castle International Centre for Lung Cancer Research, Liverpool, UK).

In a further example, SEQ ID NO:92, SEQ ID NO:93, and SEQ ID NO:94 showed differential expression in lung tumor tissues compared to normal lung tissue from the same donor as determined by microarray analysis. Samples of normal lung were compared to lung tumor from the same donor (Roy Castle International Centre for Lung Cancer Research, Liverpool, UK). The expression of SEQ ID NO:92, SEQ ID NO:93, and SEQ ID NO:94 was decreased by at least two-fold in tumor tissue as compared to the matched normal lung for seven different donors in the case of SEQ ID NO:92, and for one donor in the case of SEQ ID NO:93 and SEQ ID NO:94.

Therefore, in various embodiments, SEQ ID NO:85, and SEQ ID NO:92, SEQ ID NO:93, and SEQ ID NO:94 can each be used for one or more of the following: i) monitoring treatment of lung cancer and other cell proliferative disorders, ii) diagnostic assays for lung cancer and other cell proliferative disorders, and iii) developing therapeutics and/or other treatments for lung cancer and other cell proliferative disorders.

<u>Colon cancer</u>

In a further example, SEQ ID NO:85 showed decreased expression in sigmoid colon tumor tissue versus normal sigmoid colon tissue. Gene expression profiles were obtained by comparing normal sigmoid colon tissue from a 48-year-old female to a sigmoid colon tumor originating from a metastatic gastric sarcoma (stromal tumor) from the same donor (Huntsman Cancer Institute, Salt Lake City, UT).

In a further example, SEQ ID NO:91, SEQ ID NO:93, and SEQ ID NO:94 showed differential expression in colon tumor tissues compared to normal colon tissue from the same donor as determined by microarray analysis. Samples of normal colon were compared to colon tumor from the

same donor (Huntsman Cancer Institute, Salt Lake City, UT). The expression of SEQ ID NO:91, was decreased, and that of SEQ ID NO:93, and SEQ ID NO:94 increased, by at least two-fold in tumor tissue as compared to matched normal colon tissue.

Therefore, in various embodiments, SEQ ID NO:85, SEQ ID NO:91, SEQ ID NO:93, and SEQ ID NO:94 can each be used for one or more of the following: i) monitoring treatment of colon cancer and other cell proliferative disorders, ii) diagnostic assays for colon cancer and other cell proliferative disorders, and iii) developing therapeutics and/or other treatments for colon cancer and other cell proliferative disorders.

Ovarian cancer

In another example, SEQ ID NO:88 showed differential expression associated with ovarian cancer, as determined by microarray analysis. A normal ovary from a 79 year-old female donor was compared to an ovarian tumor from the same donor (Huntsman Cancer Institute, Salt Lake City, UT). SEQ ID NO:88 expression was increased at least two-fold in the tumor tissue as compared to the normal tissue.

In a further example, SEQ ID NO:92, SEQ ID NO:109, and SEQ ID NO:112 showed differential expression in ovary tumor versus normal ovary tissue as determined by microarray analysis. A normal ovary from a 79-year-old female donor was compared to an ovarian tumor from the same donor (Huntsman Cancer Institute, Salt Lake City, UT). The expression of SEQ ID NO:92 and SEQ ID NO:109 was increased, and the expression of SEQ ID NO:112 decreased, by at least two-fold in the ovarian tumor tissue as compared to the matched normal tissue.

Therefore, in various embodiments, SEQ ID NO:88, SEQ ID NO:92, SEQ ID NO:109, and SEQ ID NO:112 can each be used for one or more of the following: i) monitoring treatment of ovarian cancer and other cell proliferative disorders, ii) diagnostic assays for ovarian cancer and other cell proliferative disorders, and iii) developing therapeutics and/or other treatments for ovarian cancer and other cell proliferative disorders.

Osteosarcoma

In a further example, SEQ ID NO:103, SEQ ID NO:109, and SEQ ID NO:118 showed differential expression in osteosarcoma associated tissues as compared to normal osteoblasts as determined by microarray analysis. Messenger RNA from normal human osteoblasts was compared with mRNA from biopsy specimens, osteosarcoma tissues, or primary cultures or metastasized tissues. A normal osteoblast primary culture, NHOst 5488, was chosen as the reference in the initial experiments. One basic set of experiments is defined as the comparison of mRNA from biopsy specimen with that of definitive surgical specimen from the same patient. Extended study of this basic set includes mRNA from primary cell cultures of the definitive surgical specimen, muscle, or cartilage tissue from the same patient. Biopsy specimens, definitive surgical specimens, or lung metastatic tissues from different individuals were also included to reveal individual variability. Expression of SEQ ID NO:103 was increased, and expression of SEQ ID NO:109 and SEQ ID NO:118 decreased, by at least two-fold in osteosarcoma associated tissues as compared to the normal osteoblasts.

Therefore, in various embodiments, SEQ ID NO:103, SEQ ID NO:109, and SEQ ID NO:118 can be used for one or more of the following: i) monitoring treatment of osteosarcoma and other cell proliferative disorders, ii) diagnostic assays for osteosarcoma and other cell proliferative disorders, and iii) developing therapeutics and/or other treatments for osteosarcoma and other cell proliferative disorders.

Autoimmune and inflammatory disorders

In a further example, PBMCs from 3 healthy volunteer donors were stimulated *in vitro* with TNF-α for 2 hours. Treated cells were compared to untreated cells from the same donors. In a separate experiment, PBMCs from 5 healthy volunteers were incubated in the presence of pro-inflammatory cytokines (IL-1β, IL-2, IL-6, IL-8, IL-12, IL-18, IFN-γ, and TNF-α) for 2 and 4 hours. Cytokine-treated PBMCs were compared to untreated PBMCs from the same donors. In both cases, the expression of SEQ ID NO:93 and SEQ ID NO:94 was increased at least two-fold in the treated cells as compared to the untreated cells. Therefore, SEQ ID NO:93 and SEQ ID NO:94 are useful in monitoring treatment of, and diagnostic assays for, autoimmune and inflammatory disorders.

**XII. Complementary Polynucleotides**

Sequences complementary to the TRICH-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring TRICH. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are

designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of TRICH. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the TRICH-encoding transcript.

**XIII. Expression of TRICH**

Expression and purification of TRICH is achieved using bacterial or virus-based expression systems. For expression of TRICH in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express TRICH upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TRICH in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding TRICH by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus (Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945).

In most expression systems, TRICH is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Biosciences). Following purification, the GST moiety can be proteolytically cleaved from TRICH at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16).

Purified TRICH obtained by these methods can be used directly in the assays shown in Examples XVII, XVIII, and XIX, where applicable.

**XIV. Functional Assays**

TRICH function is assessed by expressing the sequences encoding TRICH at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT plasmid (Invitrogen, Carlsbad CA) and PCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 $\mu$g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 $\mu$g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994; Flow Cytometry, Oxford, New York NY).

The influence of TRICH on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding TRICH and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding TRICH and other genes of interest can be analyzed by northern analysis or microarray techniques.

### XV. Production of TRICH Specific Antibodies

TRICH substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize animals (e.g., rabbits, mice, etc.) and to produce antibodies using standard protocols.

Alternatively, the TRICH amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art (Ausubel et al., *supra*, ch. 11).

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity (Ausubel et al., *supra*). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-TRICH activity by, for example, binding the peptide or TRICH to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

### XVI. Purification of Naturally Occurring TRICH Using Specific Antibodies

Naturally occurring or recombinant TRICH is substantially purified by immunoaffinity chromatography using antibodies specific for TRICH. An immunoaffinity column is constructed by covalently coupling anti-TRICH antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Biosciences). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing TRICH are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TRICH (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TRICH binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and TRICH is collected.

### XVII. Identification of Molecules Which Interact with TRICH

Molecules which interact with TRICH may include transporter substrates, agonists or antagonists, modulatory proteins such as Gβγ proteins (Reimann, *supra*) or proteins involved in TRICH localization or clustering such as MAGUKs (Craven, *supra*). TRICH, or biologically active fragments thereof, are labeled with $^{125}$I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M.

Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled TRICH, washed, and any wells with labeled TRICH complex are assayed. Data obtained using different concentrations of TRICH are used to calculate values for the number, affinity, and association of TRICH with the candidate molecules.

Alternatively, proteins that interact with TRICH are isolated using the yeast 2-hybrid system (Fields, S. and O. Song (1989) Nature 340:245-246). TRICH, or fragments thereof, are expressed as fusion proteins with the DNA binding domain of Gal4 or lexA, and potential interacting proteins are expressed as fusion proteins with an activation domain. Interactions between the TRICH fusion protein and the TRICH interacting proteins (fusion proteins with an activation domain) reconstitute a transactivation function that is observed by expression of a reporter gene. Yeast 2-hybrid systems are commercially available, and methods for use of the yeast 2-hybrid system with ion channel proteins are discussed in Niethammer, M. and M. Sheng (1998, Meth. Enzymol. 293:104-122).

TRICH may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Potential TRICH agonists or antagonists may be tested for activation or inhibition of TRICH ion channel activity using the assays described in section XVIII.

**XVIII. Demonstration of TRICH Activity**

Ion channel activity of TRICH is demonstrated using an electrophysiological assay for ion conductance. TRICH can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding TRICH. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A second plasmid which expresses any one of a number of marker genes, such as ß-galactosidase, is co-transformed into the cells to allow rapid identification of those cells which have taken up and expressed the foreign DNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of TRICH and ß-galactosidase.

Transformed cells expressing ß-galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media under conditions that are well known in the art. Stained cells are tested for differences in membrane conductance by electrophysiological techniques that are well known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or

ß-galactosidase sequences alone, are used as controls and tested in parallel. Cells expressing TRICH will have higher anion or cation conductance relative to control cells. The contribution of TRICH to conductance can be confirmed by incubating the cells using antibodies specific for TRICH. The antibodies will bind to the extracellular side of TRICH, thereby blocking the pore in the ion channel, and the associated conductance.

Alternatively, ion channel activity of TRICH is measured as current flow across a TRICH-containing *Xenopus laevis* oocyte membrane using the two-electrode voltage-clamp technique (Ishi et al., *supra*; Jegla, T. and L. Salkoff (1997) J. Neurosci. 17:32-44). TRICH is subcloned into an appropriate *Xenopus* oocyte expression vector, such as pBF, and 0.5-5 ng of mRNA is injected into mature stage IV oocytes. Injected oocytes are incubated at 18°C for 1-5 days. Inside-out macropatches are excised into an intracellular solution containing 116 mM K-gluconate, 4 mM KCl, and 10 mM Hepes (pH 7.2). The intracellular solution is supplemented with varying concentrations of the TRICH mediator, such as cAMP, cGMP, or $Ca^{+2}$ (in the form of $CaCl_2$), where appropriate. Electrode resistance is set at 2-5 MΩ and electrodes are filled with the intracellular solution lacking mediator. Experiments are performed at room temperature from a holding potential of 0 mV. Voltage ramps (2.5 s) from -100 to 100 mV are acquired at a sampling frequency of 500 Hz. Current measured is proportional to the activity of TRICH in the assay.

Transport activity of TRICH is assayed by measuring uptake of labeled substrates into *Xenopus laevis* oocytes. Oocytes at stages V and VI are injected with TRICH mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$, 1mM $Na_2HPO_4$, 5 mM Hepes, 3.8 mM NaOH , 50$\mu$g/ml gentamycin, pH 7.8) to allow expression of TRICH. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, ions, and neurotransmitters) is initiated by adding labeled substrate (e.g. radiolabeled with $^3H$, fluorescently labeled with rhodamine, etc.) to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in $Na^+$-free medium, measuring the incorporated label, and comparing with controls. TRICH activity is proportional to the level of internalized labeled substrate.

ATPase activity associated with TRICH can be measured by hydrolysis of radiolabeled ATP-[$\gamma$-$^{32}P$], separation of the hydrolysis products by chromatographic methods, and quantitation of the recovered $^{32}P$ using a scintillation counter. The reaction mixture contains ATP-[$\gamma$-$^{32}P$] and varying amounts of TRICH in a suitable buffer incubated at 37°C for a suitable period of time. The reaction

is terminated by acid precipitation with trichloroacetic acid and then neutralized with base, and an aliquot of the reaction mixture is subjected to membrane or filter paper-based chromatography to separate the reaction products. The amount of $^{32}P$ liberated is counted in a scintillation counter. The amount of radioactivity recovered is proportional to the ATPase activity of TRICH in the assay.

Lipocalin activity of TRICH is measured by ligand fluorescence enhancement spectrofluorometry (Lin et al. (1997) Molecular Vision 3:17). Examples of ligands include retinol (Sigma, St. Louis MO) and 16-anthryloxy-palmitic acid (16-AP) (Molecular Probes Inc., Eugene OR). Ligand is dissolved in 100% ethanol and its concentration is estimated using known extinction coefficents (retinol: 46,000 A/M/cm at 325 nm; 16-AP: 8,200 A/M/cm at 361 nm). A 700 $\mu$l aliquot of 1 $\mu$M TRICH in 10 mM Tris (pH 7.5), 2 mM EDTA, and 500 mM NaCl is placed in a 1 cm path length quartz cuvette and 1 $\mu$l aliquots of ligand solution are added. Fluorescence is measured 100 seconds after each addition until readings are stable. Change in fluorescence per unit change in ligand concentration is proportional to TRICH activity.

### XIX. Identification of TRICH Agonists and Antagonists

TRICH is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK (Human Embryonic Kidney) 293. Ion channel activity of the transformed cells is measured in the presence and absence of candidate agonists or antagonists. Ion channel activity is assayed using patch clamp methods well known in the art or as described in Example XVIII. Alternatively, ion channel activity is assayed using fluorescent techniques that measure ion flux across the cell membrane (Velicelebi, G. et al. (1999) Meth. Enzymol. 294:20-47; West, M.R. and C.R. Molloy (1996) Anal. Biochem. 241:51-58). These assays may be adapted for high-throughput screening using microplates. Changes in internal ion concentration are measured using fluorescent dyes such as the $Ca^{2+}$ indicator Fluo-4 AM, sodium-sensitive dyes such as SBFI and sodium green, or the $Cl^-$ indicator MQAE (all available from Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In a more generic version of this assay, changes in membrane potential caused by ionic flux across the plasma membrane are measured using oxonyl dyes such as $DiBAC_4$ (Molecular Probes). $DiBAC_4$ equilibrates between the extracellular solution and cellular sites according to the cellular membrane potential. The dye's fluorescence intensity is 20-fold greater when bound to hydrophobic intracellular sites, allowing detection of $DiBAC_4$ entry into the cell (Gonzalez, J.E. and P.A. Negulescu (1998) Curr. Opin. Biotechnol. 9:624-631). Candidate agonists or antagonists may be selected from known ion channel agonists or antagonists, peptide libraries, or combinatorial chemical libraries.

Various modifications and variations of the described compositions, methods, and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. It will be appreciated that the invention provides novel and useful proteins, and their encoding polynucleotides, which can be used in the drug discovery process, as well as methods for using these compositions for the detection, diagnosis, and treatment of diseases and conditions. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Nor should the description of such embodiments be considered exhaustive or limit the invention to the precise forms disclosed. Furthermore, elements from one embodiment can be readily recombined with elements from one or more other embodiments. Such combinations can form a number of embodiments within the scope of the invention. It is intended that the scope of the invention be defined by the following claims and their equivalents.

# Table 1

| Incyte Project ID | Polypeptide SEQ ID NO: | Incyte Polypeptide ID | Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Incyte Full Length Clones |
|---|---|---|---|---|---|
| 7509332 | 1 | 7509332CD1 | 60 | 7509332CB1 | 90124688CA2 |
| 7509102 | 2 | 7509102CD1 | 61 | 7509102CB1 | 90134847CA2 |
| 7509132 | 3 | 7509132CD1 | 62 | 7509132CB1 | 90134560CA2 |
| 7509136 | 4 | 7509136CD1 | 63 | 7509136CB1 | 90138017CA2 |
| 7509178 | 5 | 7509178CD1 | 64 | 7509178CB1 | 90138906CA2 |
| 7509214 | 6 | 7509214CD1 | 65 | 7509214CB1 | 90138823CA2 |
| 7509244 | 7 | 7509244CD1 | 66 | 7509244CB1 | 90137849CA2 |
| 7509256 | 8 | 7509256CD1 | 67 | 7509256CB1 | 2444801CA2, 4936749CA2, 90028858CA2, 90028990CA2, 90138157CA2, 90138181CA2, 90161926CA2, 90223995CA2 |
| 7509395 | 9 | 7509395CD1 | 68 | 7509395CB1 | 90139077CA2 |
| 7503287 | 10 | 7503287CD1 | 69 | 7503287CB1 | |
| 7503320 | 11 | 7503320CD1 | 70 | 7503320CB1 | 90036682CA2, 90036790CA2, 90036818CA2 |
| 7503335 | 12 | 7503335CD1 | 71 | 7503335CB1 | |
| 7503952 | 13 | 7503952CD1 | 72 | 7503952CB1 | 90103638CA2 |
| 7504530 | 14 | 7504530CD1 | 73 | 7504530CB1 | 90017261CA2, 90219736CA2, 90219792CA2, 90219860CA2, 90220851CA2, 90220883CA2 |
| 7509303 | 15 | 7509303CD1 | 74 | 7509303CB1 | |

# Table 1

| Incyte Project ID | Polypeptide SEQ ID NO: | Incyte Polypeptide ID | Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Incyte Full Length Clones |
|---|---|---|---|---|---|
| 7509910 | 16 | 7509910CD1 | 75 | 7509910CB1 | 7049239CA2 |
| 7509982 | 17 | 7509982CD1 | 76 | 7509982CB1 | |
| 7510082 | 18 | 7510082CD1 | 77 | 7510082CB1 | |
| 7510367 | 19 | 7510367CD1 | 78 | 7510367CB1 | 90023684CA2 |
| 7510413 | 20 | 7510413CD1 | 79 | 7510413CB1 | |
| 1721303 | 21 | 1721303CD1 | 80 | 1721303CB1 | 2905327CA2, 5765782CA2 |
| 7502007 | 22 | 7502007CD1 | 81 | 7502007CB1 | 90055806CA2, 90055838CA2, 90055846CA2 |
| 7506439 | 23 | 7506439CD1 | 82 | 7506439CB1 | 90117352CA2, 90117412CA2 |
| 7509243 | 24 | 7509243CD1 | 83 | 7509243CB1 | 7616162CA2 |
| 7509404 | 25 | 7509404CD1 | 84 | 7509404CB1 | 90138209CA2, 90224278CA2 |
| 7509439 | 26 | 7509439CD1 | 85 | 7509439CB1 | 8241250CA2 |
| 7510202 | 27 | 7510202CD1 | 86 | 7510202CB1 | |
| 7510203 | 28 | 7510203CD1 | 87 | 7510203CB1 | |
| 7510208 | 29 | 7510208CD1 | 88 | 7510208CB1 | |
| 7510446 | 30 | 7510446CD1 | 89 | 7510446CB1 | 90048796CA2, 90048896CA2 |
| 7505294 | 31 | 7505294CD1 | 90 | 7505294CB1 | |
| 7505631 | 32 | 7505631CD1 | 91 | 7505631CB1 | |
| 7506561 | 33 | 7506561CD1 | 92 | 7506561CB1 | 6156076CA2 |
| 7510733 | 34 | 7510733CD1 | 93 | 7510733CB1 | |
| 7510734 | 35 | 7510734CD1 | 94 | 7510734CB1 | 90057371CA2, 95157512CA2, 95157544CA2, 95157552CA2 |

# Table 1

| Incyte Project ID | Polypeptide SEQ ID NO: | Incyte Polypeptide ID | Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Incyte Full Length Clones |
|---|---|---|---|---|---|
| 7503977 | 36 | 7503977CD1 | 95 | 7503977CB1 | |
| 7505084 | 37 | 7505084CD1 | 96 | 7505084CB1 | |
| 7506950 | 38 | 7506950CD1 | 97 | 7506950CB1 | |
| 7506951 | 39 | 7506951CD1 | 98 | 7506951CB1 | 90119019CA2 |
| 7506954 | 40 | 7506954CD1 | 99 | 7506954CB1 | 90119183CA2 |
| 7506956 | 41 | 7506956CD1 | 100 | 7506956CB1 | 90119035CA2, 90119259CA2 |
| 7506959 | 42 | 7506959CD1 | 101 | 7506959CB1 | |
| 7506960 | 43 | 7506960CD1 | 102 | 7506960CB1 | 90118991CA2, 90119051CA2, 90119110CA2, 90119118CA2, 90119127CA2, 90119174CA2, 90119218CA2, 90119251CA2, 90119258CA2, 90119266CA2, 90120004CA2 |
| 7510540 | 44 | 7510540CD1 | 103 | 7510540CB1 | 90059701CA2, 90059717CA2 |
| 7510545 | 45 | 7510545CD1 | 104 | 7510545CB1 | 90049442CA2 |
| 7510654 | 46 | 7510654CD1 | 105 | 7510654CB1 | |
| 7510660 | 47 | 7510660CD1 | 106 | 7510660CB1 | |
| 7510661 | 48 | 7510661CD1 | 107 | 7510661CB1 | |
| 7510680 | 49 | 7510680CD1 | 108 | 7510680CB1 | 90112131CA2 |
| 7505145 | 50 | 7505145CD1 | 109 | 7505145CB1 | |
| 7505162 | 51 | 7505162CD1 | 110 | 7505162CB1 | 95223082CA2 |
| 7505469 | 52 | 7505469CD1 | 111 | 7505469CB1 | |

# Table 1

| Incyte Project ID | Polypeptide SEQ ID NO: | Incyte Polypeptide ID | Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Incyte Full Length Clones |
|---|---|---|---|---|---|
| 7505475 | 53 | 7505475CD1 | 112 | 7505475CB1 | |
| 7505568 | 54 | 7505568CD1 | 113 | 7505568CB1 | 90002693CA2, 90011331CA2, 90011519CA2 |
| 7506953 | 55 | 7506953CD1 | 114 | 7506953CB1 | 90119067CA2 |
| 7510176 | 56 | 7510176CD1 | 115 | 7510176CB1 | 4730495CA2 |
| 7510541 | 57 | 7510541CD1 | 116 | 7510541CB1 | |
| 7510923 | 58 | 7510923CD1 | 117 | 7510923CB1 | |
| 7510984 | 59 | 7510984CD1 | 118 | 7510984CB1 | |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| 1 | 7509332CD1 | g182418 | 8.1E-77 | [Homo sapiens] folate-binding protein precursor |
| | | | | Elwood, P. C. |
| | | | | Molecular cloning and characterization of the human folate-binding protein cDNA from placenta and malignant tissue culture (KB) cells |
| | | | | J. Biol. Chem. 264, 14893-14901 (1989) |
| | | 606110\|FOLR1 | 6.8E-78 | [Homo sapiens] [Receptor (signalling); Small molecule-binding protein] [Unspecified membrane; Plasma membrane] Folate receptor 1 (folate receptor alpha), binds and transports folate and may play a role in neural tube morphogenesis; mutations in the corresponding gene may contribute to neural tube defects |
| | | | | Campbell, I. G. et al. |
| | | | | Folate-binding protein is a marker for ovarian cancer. |
| | | | | Cancer Res 51, 5329-38 (1991). |
| | | 582905\|Folr2 | 6.1E-77 | [Mus musculus] [Small molecule-binding protein] Folate-binding protein, high affinity, low capacity |
| | | | | Piedrahita, J. A. et al. |
| | | | | Mice lacking the folic acid-binding protein Folbp1 are defective in early embryonic development. |
| | | | | Nat Genet 23, 228-32 (1999). |
| 2 | 7509102CD1 | g1458110 | 9.7E-69 | [Homo sapiens] nicotinic acetylcholine receptor alpha2 subunit precursor |
| | | | | Elliott, K. J. et al. |
| | | | | Comparative structure of human neuronal alpha 2-alpha 7 and beta 2-beta 4 nicotinic acetylcholine receptor subunits and functional expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and beta 4 subunits |
| | | | | J. Mol. Neurosci. 7, 217-228 (1996) |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
| --- | --- | --- | --- | --- |
| | | 334652\|CHRNA2 | 8.2E-70 | [Homo sapiens] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] Cholinergic receptor nicotinic alpha polypeptide 2, a nicotinic acetylcholine-activated cation-selective channel that may play a role in signal transduction and synaptic transmission |
| | | | | Sato, K. Z. et al. |
| | | | | Diversity of mRNA expression for muscarinic acetylcholine receptor subtypes and neuronal nicotinic acetylcholine receptor subunits in human mononuclear leukocytes and leukemic cell lines. |
| | | | | Neurosci Lett 266, 17-20 (1999). |
| | | 329298\|Rn.9713 | 9.7E-41 | [Rattus norvegicus] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Cholinergic receptor nicotinic alpha polypeptide 2, a nicotinic acetylcholine-activated cation-selective channel that may play a role in signal transduction and synaptic transmission |
| | | | | Francis, M. M. et al. |
| | | | | Subtype-selective inhibition of neuronal nicotinic acetylcholine receptors by cocaine is determined by the alpha4 and beta4 subunits |
| | | | | Mol Pharmacol 58, 109-19 (2000). |
| 3 | 7509132CD1 | g183296 | 9.3E-167 | [Homo sapiens] glucose transporter |
| | | | | Buse, J. B. et al. |
| | | | | Expression and regulation of the human GLUT4/muscle-fat facilitative glucose transporter gene in transgenic mice |
| | | | | J. Biol. Chem. 267, 11673-11676 (1992) |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 338070\|SLC2A4 | 7.8E-168 | [Homo sapiens] [Active transporter, secondary; Major Facilitator Superfamily; Transporter] [Unspecified membrane; Plasma membrane] Glucose trasporter 4, a glucose transporter that translocates to the plasma membrane in response to insulin and plays a role in carbohydrate metabolism; targeted disruption of the gene for mouse Slc2a4 results in insulin resistance and diabetes |
| | | | | Oshel, K. M. et al. |
| | | | | Identification of a 30-base pair regulatory element and novel DNA binding protein that regulates the human GLUT4 promoter in transgenic mice. |
| | | | | J Biol Chem 275, 23666-73 (2000). |
| | | 430572\|Slc2a4 | 1.6E-160 | [Rattus norvegicus] [Active transporter, secondary; Major Facilitator Superfamily; Transporter] [Endosome/Endosomal vesicles; Nuclear; Endoplasmic reticulum; Cytoplasmic; Unspecified membrane; Plasma membrane] Glucose trasporter 4, a glucose transporter that translocates to the plasma membrane in response to insulin and plays a role in carbohydrate metabolism; targeted disruption of the gene for mouse Slc2a4 results in insulin resistance and diabetes |
| | | | | Kanzaki, M. et al. |
| | | | | The trimeric GTP-binding protein (G(q)/G(11)) alpha subunit is required for insulin-stimulated GLUT4 translocation in 3T3L1 adipocytes. |
| | | | | J Biol Chem 275, 7167-75 (2000). |
| 4 | 7509136CD1 | g15030222 | 1.2E-109 | [Homo sapiens] Similar to cholinergic receptor, nicotinic, beta polypeptide 1 (muscle) |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 339230\|CHRNB1 | 5.6E-110 | [Homo sapiens] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Cholinergic receptor (nicotinic) beta 1 subunit, may play an important role in neuromuscular synaptic transmission; mutations in the corresponding gene are associated with slow-channel congenital myasthenic syndromes |
| | | | | Quiram, P. A. et al. |
| | | | | Mutation causing congenital myasthenia reveals acetylcholine receptor beta/delta subunit interaction essential for assembly. |
| | | | | J Clin Invest 104, 1403-10. (1999). |
| | | 568818\|CHRNG | 1.3E-74 | [Homo sapiens] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Gamma subunit of the muscle nicotinic acetylcholine receptor, a fetal-type subunit that is replaced after birth by the epsilon subunit (CHRNE), contains antigenic epitopes that may contribute to the development of myasthenia gravis |
| | | | | Vernet-der Garabedian, B. et al. |
| | | | | Association of neonatal myasthenia gravis with antibodies against the fetal acetylcholine receptor. |
| | | | | J Clin Invest 94, 555-9 (1994). |
| 5 | 7509178CD1 | g669153 | 1.9E-159 | [Homo sapiens] acetylcholine receptor |
| | | | | Noda, M. et al. |
| | | | | Cloning and sequence analysis of calf cDNA and human genomic DNA encoding alpha-subunit precursor of muscle acetylcholine receptor |
| | | | | Nature 305, 818-823 (1983) |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 339228\|CHRNA1 | 1.6E-160 | [Homo sapiens] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Alpha subunit of the muscle nicotinic acetylcholine receptor, contains the major binding site for acetylcholine and the immunogenic site associated with autoantibodies in maysthenia gravis; mutations are associated with slow-channel myasthenic syndrome |
| | | | | Sine, S. M. et al. |
| | | | | Mutation of the acetylcholine receptor alpha subunit causes a slow- channel myasthenic syndrome by enhancing agonist binding affinity. |
| | | | | Neuron 15, 229-39 (1995). |
| | | 580847\|Chrna1 | 3.7E-154 | [Mus musculus] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Alpha subunit of the muscle nicotinic acetylcholine receptor, contains major binding site for acetylcholine and epitope for autoantibodies in experimental maysthenia gravis; mutations in human CHRNA1 are associated with slow-channel myasthenic syndrome |
| | | | | Merlie, J. P. et al. |
| | | | | Myogenin and acetylcholine receptor alpha gene promoters mediate transcriptional regulation in response to motor innervation. |
| | | | | J Biol Chem 269, 2461-7 (1994). |
| 6 | 7509214CD1 | g488420 | 2.2E-55 | [Homo sapiens] peripheral benzodiazepine receptor related protein |
| | | | | Lin, D. et al. |
| | | | | The human peripheral benzodiazepine receptor gene: cloning and characterization of alternative splicing in normal tissues and in a patient with congenital lipoid adrenal hyperplasia |
| | | | | Genomics 18, 643-650 (1993) |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 568290\|BZRP | 1.8E-56 | [Homo sapiens] [Channel (passive transporter); Receptor (protein translocation); Transporter; Receptor (signalling)] [Cytoplasmic; Mitochondrial outer membrane; Mitochondrial] Benzodiazapine receptor (peripheral), involved in steroid biosynthesis, cell proliferation, and may contribute to mitochondrial biogenesis and inhibit oxygen radical induced apoptosis; expression, nuclear location may correlate to breast tumor progression |
| | | | | Hardwick, M. et al. |
| | | | | Peripheral-type benzodiazepine receptor (PBR) in human breast cancer: correlation of breast cancer cell aggressive phenotype with PBR expression, nuclear localization, and PBR-mediated cell proliferation and nuclear transport of cholesterol. |
| | | | | Cancer Res 59, 831-42 (1999). |
| 7 | 7509244CD1 | g560155 | 1.9E-199 | [Homo sapiens] acetylcholine receptor beta-subunit preprotein |
| | | | | Beeson, D. et al. |
| | | | | Nucleotide sequence of human muscle acetylcholine receptor beta-subunit |
| | | | | Nucleic Acids Res. 17, 4391 (1989) |
| | | 339230\|CHRNB1 | 1.6E-200 | [Homo sapiens] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Cholinergic receptor (nicotinic) beta 1 subunit, may play an important role in neuromuscular synaptic transmission; mutations in the corresponding gene are associated with slow-channel congenital myasthenic syndromes |
| | o | | | Quiram, P. A. et al. (supra) |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 430536\|Chrnb1 | 1E-185 | [Rattus norvegicus] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] Cholinergic receptor (nicotinic) beta 1 subunit, expression is differentially regulated during myogenesis; mutations of the corresponding human CHRNB1 gene are associated with slow-channel congenital myasthenic syndromes |
| | | | | Witzemann, V. et al. |
| | | | | Primary structure and functional expression of the alpha-, beta-, gamma-, delta- and epsilon-subunits of the acetylcholine receptor from rat muscle. |
| | | | | Eur J Biochem 194, 437-48 (1990). |
| 8 | 7509256CD1 | g992687 | 7.7E-163 | [Homo sapiens] glycine receptor beta subunit |
| | | | | Handford, C. A. et al. |
| | | | | The human glycine receptor beta subunit: primary structure, functional characterisation and chromosomal localisation of the human and murine genes |
| | | | | Brain Res. Mol. Brain Res. 35, 211-219 (1996) |
| | | 335538\|GLRB | 6.5E-164 | [Homo sapiens] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Glycine receptor beta, a subunit of the chloride channel important for inhibitory neurotransmission |
| | | | | Milani, N. et al. |
| | | | | The human glycine receptor beta subunit gene (GLRB): structure, refined chromosomal localization, and population polymorphism. |
| | | | | Genomics 50, 341-5 (1998). |
| | | 760128\|Glrb | 1.7E-156 | [Mus musculus] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] Glycine receptor beta, a subunit of the chloride channel important for inhibitory neurotransmission; implicated in congenital myoclonus |
| | | | | Tintrup, H. et al. |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Exonic Sp1 sites are required for neural-specific expression of the glycine receptor beta subunit gene. |
| | | | | Biochem J 355, 179-87. (2001). |
| 9 | 7509395CD1 | g669153 | 6.6E-195 | [Homo sapiens] acetylcholine receptor |
| | | | | Noda, M. et al. (supra) |
| | | 339228\|CHRNA1 | 5.5E-196 | [Homo sapiens] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Alpha subunit of the muscle nicotinic acetylcholine receptor, contains the major binding site for acetylcholine and the immunogenic site associated with autoantibodies in maysthenia gravis; mutations are associated with slow-channel myasthenic syndrome |
| | | | | Sine, S. M. et al. (supra) |
| | | 580847\|Chrna1 | 1.2E-182 | [Mus musculus] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Alpha subunit of the muscle nicotinic acetylcholine receptor, contains major binding site for acetylcholine and epitope for autoantibodies in experimental maysthenia gravis; mutations in human CHRNA1 are associated with slow-channel myasthenic syndrome |
| | | | | Boulter, J. et al. |
| | | | | Isolation of a clone coding for the alpha-subunit of a mouse acetylcholine receptor. |
| | | | | J Neurosci 5, 2545-52 (1985) |
| 10 | 7503287CD1 | g1871170 | 7.8E-133 | [Homo sapiens] sodium channel 2 |
| | | | | Garcia-Anoveros, J. et al. |
| | | | | BNaC1 and BNaC2 constitute a new family of human neuronal sodium channels related to degenerins and epithelial sodium channels |
| | | | | Proc. Natl. Acad. Sci. U.S.A. 94, 1459-1464 (1997) |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 610724\|ACCN2 | 6.3E-134 | [Homo sapiens] [Channel (passive transporter); Transporter] [Plasma membrane] Amiloride-sensitive cation channel 2 (acid-sensing ion channel), a member of the DEG/ENaC superfamily of sodium channels |
| | | | | Sayegh, R. et al. |
| | | | | Glucocorticoid induction of epithelial sodium channel expression in lung and renal epithelia occurs via trans-activation of a hormone response element in the 5'-flanking region of the human epithelial sodium channel alpha subunit gene. |
| | | | | J Biol Chem 274, 12431-7 (1999). |
| | | 685845\|Accn2 | 3.1E-132 | [Rattus norvegicus] [Channel (passive transporter); Transporter] [Plasma membrane] Proton-gated cation channel (acid-sensing ion channel 1), amiloride-sensitive sodium channel that is a member of the DEG/ENaC superfamily, putatively mediates sensory perception and may define sensitivity to tarantula toxin |
| | | | | Voilley, N. et al. |
| | | | | Nonsteroid anti-inflammatory drugs inhibit both the activity and the inflammation-induced expression of acid-sensing ion channels in nociceptors. |
| | | | | J Neurosci 21, 8026-33. (2001). |
| 11 | 7503320CD1 | g2808624 | 1.7E-34 | [Homo sapiens] nicotinic acetylcholine receptor alpha7 subunit precursor |
| | | | | Groot Kormelink, P. J. et al. |
| | | | | Cloning and sequence of full-length cDNAs encoding the human neuronal nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and expression of seven nAChR subunits in the human neuroblastoma cell line SH-SY5Y and/or IMR-32 |
| | | | | FEBS Lett. 400, 309-314 (1997) |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 334660\|CHRNA7 | 1.4E-35 | [Homo sapiens] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Alpha 7 subunit of the neuronal nicotinic acetylcholine receptor, binds alpha bungarotoxin, highly permeable to calcium, may be involved in Alzheimer's disease and schizophrenia |
| | | | | Leonard, S. et al. |
| | | | | Smoking and schizophrenia: abnormal nicotinic receptor expression. |
| | | | | Eur J Pharmacol 393, 237-42 (2000). |
| | | 589947\|Chrna7 | 1.3E-31 | [Rattus norvegicus] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Alpha 7 subunit of the neuronal nicotinic acetylcholine receptor, binds alpha bungarotoxin; human CHRNA7 may be involved in Alzheimer's disease and schizophrenia |
| | | | | Dominguez del Toro, E. et al. |
| | | | | Expression of alpha 7 neuronal nicotinic receptors during postnatal development of the rat cerebellum. |
| | | | | Brain Res Dev Brain Res 98, 125-33 (1997). |
| 12 | 7503335CD1 | g1854512 | 0.0 | [Homo sapiens] ATP receptor |
| | | | | Rassendren, F. et al. |
| | | | | The permeabilizing ATP receptor (P2X7): Cloning and expression of human cDNA |
| | | | | J Biol Chem 272, 5482-6 (1997). |
| | | 336736\|P2RX7 | 0.0 | [Homo sapiens] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] Purinergic receptor P2X (channel-7), ATP-gated cation channel capable of forming macropores permeable to large molecules, mediates macrophage lysis, IL-1beta (IL1B) release and cell fusion |
| | | | | Humphreys, B. D. et al. |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID - | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Stress-activated protein kinase/JNK activation and apoptotic induction by the macrophage P2X7 nucleotide receptor. |
| | | | | J Biol Chem 275, 26792-8 (2000). |
| | | 609795\|P2rx7 | 2.1E-251 | [Rattus norvegicus] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] Purinergic receptor P2X (channel-7), ATP-gated cation channel capable of forming macropores permeable to large molecules, mediates macrophage lysis and may play a role in fast synaptic transmission |
| | | | | Boue-Grabot, E. et al. |
| | | | | A protein kinase C site highly conserved in P2X subunits controls the desensitization kinetics of P2X(2) ATP-gated channels. |
| | | | | J Biol Chem 275, 10190-5 (2000). |
| 13 | 7503952CD1 | g4218949 | 2.1E-123 | [Homo sapiens] 5-hydroxytryptamine 3 receptor B subunit precursor |
| | | | | Davies, P. A. et al. |
| | | | | The 5-HT3B subunit is a major determinant of serotonin-receptor function |
| | | | | Nature 397, 359-363 (1999) |
| 13 | 7503952CD1 | 343014\|HTR3B | 1.7E-124 | [Homo sapiens] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] 5-hydroxytryptamine 3B (serotonin) receptor subunit, ligand-gated cation channel subunit that is coexpressed in brain with 5-HT 3A receptor subunit (HTR3A), forms heteromers with HTR3A that exhibit serotonin-induced single-channel conductance |
| | | | | Dang, H. et al. |
| | | | | Probing the role of a conserved M1 proline residue in 5-hydroxytryptamine(3) receptor gating. |
| | | | | Mol Pharmacol 57, 1114-22 (2000). |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| 13 | 7503952CD1 | 611280\|Htr3b | 1.1E-97 | [Mus musculus] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] 5-hydroxytryptamine 3B (serotonin) receptor subunit, putative ligand-gated cation channel subunit that is coexpressed with the 5-HT 3A receptor subunit (Htr3a) in several cell lines; predicted to form serotonin-responsive heteromers with Htr3a in neurons |
| | | | | Hanna, M. C. et al. |
| | | | | Evidence for expression of heteromeric serotonin 5-HT(3) receptors in rodents |
| | | | | J Neurochem 75, 240-7 (2000). |
| 14 | 7504530CD1 | g2317274 | 3.4E-132 | [Homo sapiens] aquaporin adipose |
| | | | | Kuriyama, H. et al. |
| | | | | Molecular cloning and expression of a novel human aquaporin from adipose tissue with glycerol permeability |
| | | | | Biochem. Biophys. Res. Commun. 241, 53-58 (1997) |
| | | 339834\|AQP7 | 2.7E-133 | [Homo sapiens] [Channel (passive transporter); Transporter] [Plasma membrane] Aquaporin 7, a member of the aquaporin family of water channels, facilitates transport of water and glycerol, may regulate energy balance by facilitating adipocyte glycerol release, plays a likely role in cell volume control and pinocytosis |
| | | | | Kishida, K. et al. |
| | | | | Aquaporin adipose, a putative glycerol channel in adipocytes. |
| | | | | J Biol Chem 275, 20896-902 (2000). |
| | | 583605\|Aqp7 | 2.2E-101 | [Mus musculus] [Channel (passive transporter); Transporter] [Cytoplasmic; Plasma membrane] Aquaporin 7, a member of the aquaporin family of water channels, facilitates transport of water and glycerol, may regulate glucose homeostasis by facilitating adipocyte glycerol release, may play a role in renal water resorption |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Nejsum, L. N. et al. |
| | | | | Localization of aquaporin-7 in rat and mouse kidney using RT-PCR, immunoblotting, and immunocytochemistry |
| | | | | Biochem Biophys Res Commun 277, 164-70 (2000). |
| 15 | 7509303CD1 | g4731109 | 1.1E-124 | [Homo sapiens] epithelial sodium channel alpha-subunit |
| | | | | Chow, Y. H. et al. |
| | | | | Hormonal regulation and genomic organization of the human amiloride-sensitive epithelial sodium channel alpha-subunit gene |
| | | | | Pediatr. Res. 46, 208-214 (1999) |
| | | 337892\|SCNN1A | 9.2E-126 | [Homo sapiens] [Channel (passive transporter); Transporter] [Plasma membrane; Unspecified membrane] Sodium channel (nonvoltage-gated) channel 1 alpha subunit, a component of an amiloride-sensitive channel, may function in fluid and ion homeostasis; mutations in the corresponding gene are linked to pseudohypoaldosteronism type 1 and salt malabsorption |
| | | | | Harvey, K. F. et al. |
| | | | | The Nedd4-like Protein KIAA0439 Is a Potential Regulator of the Epithelial Sodium Channel. |
| | | | | J Biol Chem 276, 8597-8601. (2001). |
| | | 711406\|Scnn1a | 5. E-102 | [Rattus norvegicus] [Channel (passive transporter); Transporter] [Plasma membrane; Unspecified membrane] Sodium channel (nonvoltage-gated) 1 alpha subunit, a component of an amiloride-sensitive channel, functions in electrolyte homeostasis; mutations in human SCNN1A are linked to pseudohypoaldosteronism type 1 characterized by salt malabsorption |
| | | | | Li, X. J. et al. |
| | | | | Alternatively spliced forms of the alpha subunit of the epithelial sodium channel: distinct sites for amiloride binding and channel pore. |
| | | | | Mol Pharmacol 47, 1133-40 (1995). |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| 16 | 7509910CD1 | g2597927 | 4.4E-194 | [Homo sapiens] P2X7 receptor |
| | | | | Buell, G. N. et al. |
| | | | | Gene structure and chromosomal localization of the human P2X7 receptor |
| | | | | Receptors Channels 5, 347-54 (1998). |
| | | 336736\|P2RX7 | 1.2E-194 | [Homo sapiens] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] Purinergic receptor P2X (channel-7), ATP-gated cation channel capable of forming macropores permeable to large molecules, mediates macrophage lysis, IL-1beta (IL1B) release and cell fusion |
| | | | | Rassendren, F. et al. (supra) |
| | | | | The permeabilizing ATP receptor, P2X7. Cloning and expression of a human cDNA. |
| | | | | J Biol Chem 272, 5482-6 (1997). |
| | | 609795\|P2rx7 | 7.5E-163 | [Rattus norvegicus] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] Purinergic receptor P2X (channel-7), ATP-gated cation channel capable of forming macropores permeable to large molecules, mediates macrophage lysis and may play a role in fast synaptic transmission |
| | | | | Boue-Grabot, E. et al. (supra) |
| 17 | 7509982CD1 | g17223622 | 0.0 | [Homo sapiens] ATP-binding cassette A6 |
| | | 568162\|ABCA8 | 0.0 | [Homo sapiens] [ATP-binding cassette; Active transporter, primary; Hydrolase; Transporter; ATPase] [Unspecified membrane; Plasma membrane] ATP-binding cassette subfamily A member 8, a putative transporter |
| | | | | Kaminski, W. E. et al. |
| | | | | ABCA6, a novel a subclass ABC transporter. |
| | | | | Biochem Biophys Res Commun 285, 1295-301. (2001). |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 333996\|ABCA3 | 6.1E-126 | [Homo sapiens] [ATP-binding cassette; Active transporter, primary; Hydrolase; Transporter; ATPase] [Unspecified membrane] ATP-binding cassette subfamily A member 3 (ATP-binding cassette transporter C), a putative transporter that is a member of the ATP-binding cassette superfamily and may have a role in development of resistance to xenobiotics |
| | | | | Klucken, J. et al. |
| | | | | ABCG1 (ABC8), the human homolog of the Drosophila white gene, is a regulator of macrophage cholesterol and phospholipid transport. |
| | | | | Proc Natl Acad Sci U S A 97, 817-22 (2000). |
| 18 | 7510082CD1 | g7415511 | 0.0 | [Homo sapiens] peptide transporter 3 |
| | | 662681\|Ci1 | 1.8E-248 | [Mus musculus] Protein induced by 8-bromo-cyclicAMP in RAW264 macrophages |
| | | | | Takahashi, Y. et al. |
| | | | | Identification of cAMP analogue inducible genes in RAW264 macrophages. |
| | | | | Biochim Biophys Acta 1492, 385-94 (2000). |
| | | 331098\|Rn.10770 | 3.3E-144 | [Rattus norvegicus] [Active transporter, secondary; Transporter] [Unspecified membrane] Peptide-histidine transporter 1, a proton-dependent high-affinity histidine transporter that also transports peptides, may also be involved in the uptake of nutritional peptides, neuromodulators, and degraded neuropeptides |
| | | | | Yamashita, T. et al. |
| | | | | Cloning and functional expression of a brain peptide/histidine transporter. |
| | | | | J Biol Chem 272, 10205-11 (1997). |
| 19 | 7510367CD1 | g11545417 | 1.1E-21 | [Homo sapiens] folate transporter/carrier |
| | | | | Titus, S. A. et al. |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Retrovirally mediated complementation of the glyB phenotype. Cloning of a human gene encoding the carrier for entry of folates into mitochondria |
| | | | | J. Biol. Chem. 275, 36811-36817 (2000) |
| | | 700836\|LOC81034 | 9.2E-23 | [Homo sapiens] Mitochondrial folate transporter |
| 20 | 7510413CD1 | g473236 | 2.4E-106 | [Homo sapiens] folate receptor FRGAMMA |
| | | | | Shen, F. et al. |
| | | | | Identification of a novel folate receptor, a truncated receptor, and receptor type beta in hematopoietic cells: cDNA cloning, expression, immunoreactivity, and tissue specificity |
| | | | | Biochemistry 33, 1209-1215 (1994) |
| | | 335364\|FOLR3 | 2.0E-107 | [Homo sapiens] [Receptor (signalling); Small molecule-binding protein] [Unspecified membrane] Folate receptor 3 (gamma), one of a family of folate receptors that includes FOLR1 and FOLR2, binds folic acid, primarily a secreted protein due to lack of an efficient signal for glycosylphosphatidylinositol anchor modification |
| | | | | Shen, F. et al. |
| | | | | Structure and regulation of a polymorphic gene encoding folate receptor type gamma/gamma'. |
| | | | | Nucleic Acids Res 26, 2132-42 (1998). |
| | | 335362\|FOLR2 | 5.9E-83 | [Homo sapiens] [Small molecule-binding protein] [Unspecified membrane] Placental folate-binding protein (folate receptor beta) |
| | | | | Ross, J. F. et al. |
| | | | | Folate receptor type beta is a neutrophilic lineage marker and is differentially expressed in myeloid leukemia. |
| | | | | Cancer 85, 348-57. (1999). |
| 21 | 1721303CD1 | g3335128 | 1.7E-20 | [Homo sapiens] F1Fo-ATPase synthase f subunit |
| | | | | Mao, M. et al. |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Identification of genes expressed in human CD34(+) hematopoietic stem/progenitor cells by expressed sequence tags and efficient full-length cDNA cloning |
| | | | | Proc. Natl. Acad. Sci. U.S.A. 95, 8175-8180 (1998) |
| 22 | 7502007CD1 | g2317274 | 1.4E-128 | [Homo sapiens] aquaporin adipose |
| | | | | Kuriyama, H. et al. (supra) |
| 23 | 7506439CD1 | g681914 | 3.1E-74 | [Homo sapiens] serotonin 5-HT3 receptor |
| | | | | Miyake, A. et al. |
| | | | | Molecular cloning of human 5-hydroxytryptamine3 receptor: heterogeneity in distribution and function among species |
| | | | | Mol. Pharmacol. 48, 407-416 (1995) |
| | | 335904\|HTR3A | 2.5E-75 | [Homo sapiens] [Channel (passive transporter); Receptor (signalling); Transporter] [Plasma membrane] 5-hydroxytryptamine receptor 3A, a serotonin receptor that is a ligand-gated ion channel, mediates a variety of physiological effects in the central and peripheral nervous system |
| | | | | Bedford, F. K. et al. |
| | | | | Neuronal expression of the 5HT3 serotonin receptor gene requires nuclear factor 1 complexes. |
| | | | | J Neurosci 18, 6186-94 (1998). |
| | | 587085\|Htr3a | 1.9E-61 | [Mus musculus] [Channel (passive transporter); Transporter; Receptor (signalling)] [Golgi; Endoplasmic reticulum; Cytoplasmic; Plasma membrane] 5-hydroxytryptamine receptor 3A, a serotonin receptor that is a ligand-gated ion channel, mediates a variety of physiological effects in the central and peripheral nervous system |
| | | | | Miquel, M. C. et al. |
| | | | | Developmental changes in the differential expression of two serotonin 5- HT3 receptor splice variants in the rat. |
| | | | | J Neurochem 65, 475-83 (1995). |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| 24 | 7509243CD1 | g13926108 | 4.6E-57 | [Homo sapiens] 2P domain potassium channel Talk-1 |
| | | | | Girard, C. et al. |
| | | | | Genomic and functional characteristics of novel human pancreatic 2P domain K+ channels |
| | | | | Biochem Biophys Res Commun 282, 249-56. (2001). |
| | | 716701\|KCNK16 | 3.7E-58 | [Homo sapiens] Potassium channel subfamily K member 16 (Twik-related alkaline pH activated K+ channel 1), a subunit of a pancreatic 2P domain background potassium channel that is open at all membrane potentials and is activated at alkaline pH |
| | | 743114\|KCNK2 | 9.3E-18 | [Homo sapiens] [Channel (passive transporter); Transporter] [Plasma membrane] Potassium channel subfamily K member 2, outwardly rectifying K+ channel, activated by volatile anesthetics, inhibited by activated protein kinases A and C |
| | | | | Medhurst, A. D. et al. |
| | | | | Distribution analysis of human two pore domain potassium channels in tissues of the central nervous system and periphery. |
| | | | | Brain Res Mol Brain Res 86, 101-114. (2001). |
| 25 | 7509404CD1 | g992687 | 4.2E-16 | [Homo sapiens] glycine receptor beta subunit |
| | | | | Handford, C. A. et al. (supra) |
| | | 335538\|GLRB | 3.4E-17 | [Homo sapiens] [Channel (passive transporter); Transporter; Receptor (signalling)] [Plasma membrane] Glycine receptor beta, a subunit of the chloride channel important for inhibitory neurotransmission |
| | | | | Handford, C. A. et al. (supra) |
| | | | | Milani, N. et al. (supra) |
| 26 | 7509439CD1 | g12654223 | 3.8E-69 | [Homo sapiens] ATP synthase, H+ transporting, mitochondrial F1 complex, gamma polypeptide 1 |
| 27 | 7510202CD1 | g17223624 | 0.0 | [Homo sapiens] ATP-binding cassette A9 |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 568162\|ABCA8 | 0.0 | [Homo sapiens] [ATP-binding cassette; Active transporter, primary; Hydrolase; Transporter; ATPase] [Unspecified membrane; Plasma membrane] ATP-binding cassette subfamily A member 8, a putative transporter |
| | | | | Kaminski, W. E. et al. (supra) |
| | | 333996\|ABCA3 | 5.1E-95 | [Homo sapiens] [ATP-binding cassette; Active transporter, primary; Hydrolase; Transporter; ATPase] [Unspecified membrane] ATP-binding cassette subfamily A member 3 (ATP-binding cassette transporter C), a putative transporter that is a member of the ATP-binding cassette superfamily and may have a role in development of resistance to xenobiotics |
| | | | | Klucken, J. et al. (supra) |
| 28 | 7510203CD1 | g15130910 | 8E-42 | [Canis familiaris] multidrug resistance protein 2 |
| | | | | Conrad, S. et al. |
| | | | | Sequencing and tissue distribution of the canine MRP2 gene compared with MRP1 and MDR1 |
| | | | | Toxicology. 156, 81-91 (2001) |
| | | 626794\|LOC64052 | 4.9E-43 | [Rattus norvegicus] [ATP-binding cassette; Active transporter, primary; Hydrolase; Transporter; ATPase] [Plasma membrane] Multidrug resistance protein, an ATP-binding cassette transporter that acts as a multidrug efflux pump |
| | | | | Saito, T. et al. |
| | | | | Expression of multidrug resistance protein 1 (MRP1) in the rat cochlea with special reference to the blood-inner ear barrier. |
| | | | | Brain Res 895, 253-7. (2001). |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 1103\|YCF1 | 6.9E-42 | [Saccharomyces cerevisiae] [ATP-binding cassette; Active transporter, primary; Hydrolase; Transporter; ATPase] [Lysosome/vacuole; Unspecified membrane] Vacuolar glutathione S-conjugate transporter, member of the ATP-binding cassette (ABC) superfamily |
| | | | | Balzi, E. et al. |
| | | | | Yeast multidrug resistance: the PDR network. |
| | | | | J Bioenerg Biomembr 27, 71-6 (1995). |
| 29 | 7510208CD1 | g9957467 | 0.0 | [Homo sapiens] ATP-binding cassette sub-family A member 2 |
| | | | | Vulevic, B. et al. |
| | | | | Cloning and characterization of human adenosine 5'-triphosphate-binding cassette, sub-family A, transporter 2 (ABCA2) |
| | | | | Cancer Res. 61, 3339-3347 (2001) |
| 30 | 7510446CD1 | g398161 | 2.2E-53 | [Homo sapiens] human ClC-1 muscle chloride channel |
| | | | | Steinmeyer, K. et al. |
| | | | | Multimeric structure of ClC-1 chloride channel revealed by mutations in dominant myotonia congenita (Thomsen) |
| | | | | EMBO J. 13, 737-743 (1994) |
| | | 334688\|CLCN1 | 1.8E-54 | [Homo sapiens] [Channel (passive transporter); Transporter] [Plasma membrane] Chloride channel 1 (skeletal muscle), transports chloride which affects muscle contraction; mutations in human CLCN1 and mouse Clcn1 genes are associated with Becker disease and Thomsen disease, both characterized by muscle membrane hyperexcitability |
| | | | | Zhang, J. et al. |
| | | | | Mechanism of inverted activation of ClC-1 channels caused by a novel myotonia congenita mutation. |
| | | | | J Biol Chem 275, 2999-3005. (2000). |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 589953\|Clcn1 | 4.7E-45 | [Rattus norvegicus] [Channel (passive transporter); Transporter] [Plasma membrane] Chloride channel 1 (skeletal muscle), transports chloride which affects muscle contraction; mutations in human CLCN1 and mouse Clcn1 genes are associated with Becker disease and Thomsen disease, both characterized by muscle membrane hyperexcitability |
| | | | | Enz, R. et al. |
| | | | | Expression of the voltage-gated chloride channel ClC-2 in rod bipolar cells of the rat retina. |
| | | | | J Neurosci 19, 9841-7 (1999). |
| 31 | 7505294CD1 | g7576452 | 2.0E-148 | [Homo sapiens] potent brain type organic ion transporter |
| | | 476069\|LOC51310 | 1.7E-149 | [Homo sapiens][Transporter][Plasma membrane; Unspecified membrane] Member of the sugar transporter family, has low similarity to rat 1-Oct, which is an organic cation transporter with broad specificity, and which is likely involved in drug elimination in kidney and liver |
| | | 430266\|Slc22a3 | 1.6E-22 | [Mus musculus][Active transporter, secondary; Major Facilitator Superfamily; Transporter][Unspecified membrane; Plasma membrane] Solute carrier family 22 member 3 (extraneuronal monoamine transporter), regulates monoamine transport in the heart and placenta |
| | | | | Kekuda, R. et al. |
| | | | | Cloning and functional characterization of a potential-sensitive, polyspecific organic cation transporter (OCT3) most abundantly expressed in placenta. |
| | | | | J Biol Chem 273, 15971-9 (1998). |
| | | | | Impaired activity of the extraneuronal monoamine transporter system known as uptake-2 in Orct3/Slc22a3-deficient mice. |
| | | | | Mol Cell Biol 21, 4188-96. (2001). |
| 32 | 7505631CD1 | 598972\|FLJ11274 | 6.7E-132 | [Homo sapiens] Protein with weak similarity to S. cerevisiae Atx2p, which is a manganese-trafficking protein |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| 33 | 7506561CD1 | g473236 | 1.1E-46 | [Homo sapiens] folate receptor FRGAMMA |
| | | | | Shen, F. et al. |
| | | | | Identification of a novel folate receptor, a truncated receptor, and receptor type beta in hematopoietic cells: cDNA cloning, expression, immunoreactivity, and tissue specificity |
| | | | | Biochemistry 33, 1209-1215 (1994) |
| | | 335364\|FOLR3 | 9.3E-48 | [Homo sapiens][Receptor (signaling); Small molecule-binding protein][Unspecified membrane] Folate receptor 3 (gamma), one of a family of folate receptors that includes FOLR1 and FOLR2, binds folic acid, primarily a secreted protein (unlike FOLR1 and FOLR2) and may be a potential drug target in CML and AML leukemias |
| | | | | Shen, F. et al. (supra) |
| | | | | Wang, H. et al. |
| | | | | Structure and regulation of a polymorphic gene encoding folate receptor type gamma/gamma'. |
| | | | | Nucleic Acids Res 26, 2132-42 (1998). |
| | | 335362\|FOLR2 | 1.6E-34 | [Homo sapiens][Small molecule-binding protein][Unspecified membrane] Placental folate-binding protein (folate receptor beta) |
| | | | | Ross, J. F. et al. |
| | | | | Folate receptor type beta is a neutrophilic lineage marker and is differentially expressed in myeloid leukemia. |
| | | | | Cancer 85, 348-57. (1999). |
| 34 | 7510733CD1 | g2887407 | 2.6E-126 | [Homo sapiens] aquaporin 9 |
| | | | | Ishibashi, K. et al. |
| | | | | Cloning and functional expression of a new aquaporin (AQP9 abundantly expressed in the peripheral leukocytes permeable to water and urea, but not to glycerol |
| | | | | Biochem. Biophys. Res. Commun. 244, 268-274 (1998) |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 613389\|AQP9 | 2.2E-127 | [Homo sapiens][Channel (passive transporter);Transporter][Plasma membrane] Aquaporin 9, member of the aquaporin channel family, mediates the transport of water and urea, expressed predominantly in leukocytes where it may play a role in immunological function |
| | | | | Ishibashi, K. et al. (supra) |
| | | 662851\|Aqp9 | 1.1E-95 | [Rattus norvegicus][Channel (passive transporter); Transporter][Plasma membrane] Neutral solute channel aquaporin 9, member of the aquaporin channel family, functions as a neutral solute channel with broad selectivity, mediates the transport of water and many non-charged solutes including carbamides, polyols, purines, and pyrimidines |
| | | | | Elkjaer, M. et al. |
| | | | | Immunolocalization of AQP9 in liver, epididymis, testis, spleen, and brain |
| | | | | Biochem Biophys Res Commun 276, 1118-28 (2000). |
| | | | | Pastor-Soler, N. et al. |
| | | | | Aquaporin 9 expression along the male reproductive tract. |
| | | | | Biol Reprod 65, 384-93. (2001). |
| 35 | 7510734CD1 | g2887407 | 1.4E-83 | [Homo sapiens] aquaporin 9 |
| | | | | Ishibashi, K. et al. (supra) |
| | | 613389\|AQP9 | 1.2E-84 | [Homo sapiens][Channel (passive transporter); Transporter][Plasma membrane] Aquaporin 9, member of the aquaporin channel family, mediates the transport of water and urea, expressed predominantly in leukocytes where it may play a role in immunological function |
| | | | | Ishibashi, K. et al. (supra) |
| | | 662851\|Aqp9 | 3.0E-65 | [Rattus norvegicus][Channel (passive transporter); Transporter][Plasma membrane] Neutral solute channel aquaporin 9, member of the aquaporin channel family, functions as a neutral solute channel with broad selectivity, mediates the transport of water and many non-charged solutes including carbamides, polyols, purines, and pyrimidines |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Elkjaer, M. et al. (supra) |
| | | | | Pastor-Soler, N. et al. (supra) |
| 36 | 7503977CD1 | g15617229 | 6.8E-109 | [Homo sapiens] TRP-related cation influx channel |
| | | | | Xu, X. Z. S. et al. |
| | | | | Regulation of melastatin, a TRP-related protein, through interaction with a cytoplasmic isoform |
| | | | | Proc. Natl. Acad. Sci. U.S.A. 98, 10692-10697 (2001) |
| 37 | 7505084CD1 | g17223724 | 2.2E-197 | [Homo sapiens] sodium/glucose cotransporter KST1 |
| | | | | Roll, P. et al. |
| | | | | New human sodium/glucose cotransporter gene (KST1): identification, characterization, and mutation analysis in ICCA (infantile convulsions and choreoathetosis) and BFIC (benign familial infantile convulsions) families |
| | | | | Gene 285, 141-148 (2002) |
| | | 762539\|RKST1 | 1.9E-198 | [Homo sapiens] Protein with high similarity to sodium-glucose cotransporter 1 (human SLC5A1), which is a high affinity glucose transporter associated with glucose-galactose malabsorption syndrome, member of the sodium:solute symporter family of membrane transporters |
| | | 590623\|Slc5a1 | 8.1E-120 | [Rattus norvegicus][Active transporter, secondary; Transporter][Unspecified membrane; Plasma membrane] Sodium-glucose cotransporter 1, a high affinity glucose transporter that is inhibited by phlorizin; mutation in human SLC5A1 is associated with glucose-galactose malabsorption syndrome |
| | | | | You, G. et al. |
| | | | | Molecular characteristics of Na(+)-coupled glucose transporters in adult and embryonic rat kidney. |
| | | | | J Biol Chem 270, 29365-71 (1995). |
| | | | | Corpe, C. et al. |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Effects of type-2 diabetes and troglitazone on the expression patterns of small intestinal sugar transporters and ppar-gamma in the zucker diabetic fatty rat. |
| | | | | Digestion 63, 116-23. (2001). |
| 38 | 7506950CD1 | g386422 | 1.7E-82 | [Homo sapiens] gamma-aminobutyric acidA receptor alpha 2 subunit; GABAA receptor alpha 2 |
| | | | | Hadingham, K. L. et al. |
| | | | | Cloning of cDNA sequences encoding human alpha 2 and alpha 3 gamma-aminobutyric acidA receptor subunits and characterization of the benzodiazepine pharmacology of recombinant alpha 1-, alpha 2-, alpha 3-, and alpha 5-containing human gamma-aminobutyric acidA receptors |
| | | | | Mol. Pharmacol. 43, 970-975 (1993) |
| | | 339368\|GABRA2 | 1.4E-83 | [Homo sapiens][Channel (passive transporter); Receptor (signaling); Transporter][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Hadingham, K. L. et al. (supra) |
| | | | | Loup, F. et al. |
| | | | | Selective alterations in GABAA receptor subtypes in human temporal lobe epilepsy |
| | | | | J Neurosci 20, 5401-19 (2000). |
| | | 582959\|Gabra2 | 1.3E-78 | [Mus musculus][Channel (passive transporter); Transporter; Receptor (signaling)][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a putative chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Sibille, E. et al. |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Genetic inactivation of the Serotonin(1A) receptor in mice results in downregulation of major GABA(A) receptor alpha subunits, reduction of GABA(A) receptor binding, and benzodiazepine-resistant anxiety. |
| | | | | J Neurosci 20, 2758-65 (2000). |
| | | | | Bouilleret, V. et al. |
| | | | | Early loss of interneurons and delayed subunit-specific changes in GABA(A)-receptor expression in a mouse model of mesial temporal lobe epilepsy |
| | | | | Hippocampus 10, 305-24 (2000). |
| 39 | 7506951CD1 | g386422 | 3.4E-153 | [Homo sapiens] gamma-aminobutyric acidA receptor alpha 2 subunit; GABAA receptor alpha 2 |
| | | | | Hadingham, K. L. et al. (supra) |
| | | 339368\|GABRA2 | 2.9E-154 | [Homo sapiens][Channel (passive transporter); Receptor (signaling); Transporter][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Hadingham, K. L. et al. (supra) |
| | | | | Loup, F. et al. (supra) |
| | | 582959\|Gabra2 | 2.8E-149 | [Mus musculus][Channel (passive transporter); Transporter; Receptor (signaling)][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a putative chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Sibille, E. et al. (supra) |
| | | | | Bouilleret, V. et al. (supra) |
| 40 | 7506954CD1 | g386422 | 7.9E-28 | [Homo sapiens] gamma-aminobutyric acidA receptor alpha 2 subunit; GABAA receptor alpha 2 |
| | | | | Hadingham, K. L. et al. (supra) |

# Table 2

<table>
<tr><th>Polypeptide SEQ ID NO:</th><th>Incyte Polypeptide ID</th><th>GenBank ID NO: or PROTEOME ID NO:</th><th>Probability Score</th><th>Annotation</th></tr>
<tr><td></td><td></td><td>339368|GABRA2</td><td>6.7E-29</td><td>[Homo sapiens][Channel (passive transporter); Receptor (signaling); Transporter][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a chloride channel that is the major inhibitory neurotransmitter receptor in the brain</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Hadingham, K. L. et al. (supra)</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Loup, F. et al. (supra)</td></tr>
<tr><td></td><td></td><td>582959|Gabra2</td><td>1.7E-23</td><td>[Mus musculus][Channel (passive transporter); Transporter; Receptor (signaling)][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a putative chloride channel that is the major inhibitory neurotransmitter receptor in the brain</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Sibille, E. et al. (supra)</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Bouilleret, V. et al. (supra)</td></tr>
<tr><td>41</td><td>7506956CD1</td><td>g386422</td><td>5.4E-215</td><td>[Homo sapiens] gamma-aminobutyric acidA receptor alpha 2 subunit; GABAA receptor alpha 2</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Hadingham, K. L. et al. (supra)</td></tr>
<tr><td></td><td></td><td>339368|GABRA2</td><td>4.6E-216</td><td>[Homo sapiens][Channel (passive transporter); Receptor (signaling); Transporter][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a chloride channel that is the major inhibitory neurotransmitter receptor in the brain</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Hadingham, K. L. et al. (supra)</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Loup, F. et al. (supra)</td></tr>
<tr><td></td><td></td><td>582959|Gabra2</td><td>4.3E-211</td><td>[Mus musculus][Channel (passive transporter); Transporter; Receptor (signaling)][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a putative chloride channel that is the major inhibitory neurotransmitter receptor in the brain</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Sibille, E. et al. (supra)</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Bouilleret, V. et al. (supra)</td></tr>
<tr><td>42</td><td>7506959CD1</td><td>g369</td><td>6.1E-214</td><td>[Bos taurus] GABA-A receptor alpha-2 precursor</td></tr>
</table>

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Schofield, P. R. et al. |
| | | | | Structural and functional basis for GABAA receptor heterogeneity |
| | | | | Nature 335, 76-79 (1988) |
| | | 339368\|GABRA2 | 1.7E-216 | [Homo sapiens][Channel (passive transporter); Receptor (signaling); Transporter][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Hadingham, K. L. et al. (supra) |
| | | | | Loup, F. et al. (supra) |
| | | 582959\|Gabra2 | 1.6E-211 | [Mus musculus][Channel (passive transporter); Transporter; Receptor (signaling)][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a putative chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Sibille, E. et al. (supra) |
| | | | | Bouilleret, V. et al. (supra) |
| 43 | 7506960CD1 | g386422 | 1.3E-27 | [Homo sapiens] gamma-aminobutyric acidA receptor alpha 2 subunit; GABAA receptor alpha 2 |
| | | | | Hadingham, K. L. et al. (supra) |
| | | 339368\|GABRA2 | 1.1E-28 | [Homo sapiens][Channel (passive transporter); Receptor (signaling); Transporter][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Hadingham, K. L. et al. (supra) |
| | | | | Loup, F. et al. (supra) |
| | | 582959\|Gabra2 | 2.8E-23 | [Mus musculus][Channel (passive transporter); Transporter; Receptor (signaling)][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a putative chloride channel that is the major inhibitory neurotransmitter receptor in the brain |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Sibille, E. et al. (supra) |
| | | | | Bouilleret, V. et al. (supra) |
| 44 | 7510540CD1 | g307125 | 6.3E-13 | [Homo sapiens] glucose transporter-like protein |
| | | | | Fukumoto, H. et al. |
| | | | | Sequence, tissue distribution, and chromosomal localization of mRNA encoding a human glucose transporter-like protein |
| | | | | Proc. Natl. Acad. Sci. U.S.A. 85, 5434-5438 (1988) |
| | | 339590\|SLC2A2 | 5.4E-14 | [Homo sapiens][Active transporter, secondary; Major Facilitator Superfamily; Transporter][Unspecified membrane; Plasma membrane] Facilitative glucose transporter 2, a low-affinity, high-capacity glucose transporter; mutations in the gene may cause Fanconi-Bickel syndrome and may be associated with pathogenesis of non-insulin-dependent diabetes |
| | | | | Santer, R. et al. |
| | | | | Mutations in GLUT2, the gene for the liver-type glucose transporter, in patients with Fanconi-Bickel syndrome |
| | | | | [published erratum appears in Nat Genet 1998 Mar;18(3):298] |
| | | 704295\|Slc2a2 | 2.1E-12 | [Mus musculus][Active transporter, secondary;Major Facilitator Superfamily;Transporter][Plasma membrane] Facilitative glucose transporter 2, a putative low-affinity, high-capacity glucose transporter, may act as a glucose sensor; mutations in human SLC2A2 may cause Fanconi-Bickel syndrome and may be associated with non-insulin-dependent diabetes |
| | | | | Thorens, B. et al. |
| | | | | Transgenic reexpression of GLUT1 or GLUT2 in pancreatic beta cells rescues GLUT2-null mice from early death and restores normal glucose-stimulated insulin secretion. |
| | | | | J Biol Chem 275, 23751-8 (2000). |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| 45 | 7510545CD1 | g306850 | 8.9E-42 | [Homo sapiens] HK1 |
| | | | | Lytton, J. et al. |
| | | | | Molecular cloning of cDNAs from human kidney coding for two alternatively spliced products of the cardiac Ca2+-ATPase gene |
| | | | | J. Biol. Chem. 263, 15024-15031 (1988) |
| | | 334260\|ATP2A2 | 6.4E-43 | [Homo sapiens][Active transporter, primary; Hydrolase; Transporter; ATPase][Endoplasmic reticulum; Microsomal fraction; Cytoplasmic; Unspecified membrane; Plasma membrane] Sarcoplasmic reticulum Ca(2+)-ATPase 2, slow twich muscle, cardiac and nonmuscle form, pumps calcium from cytoplasm into ER; reduced activity in the heart is implicated in dilated cardiomyopathy and gene mutations are associated with Darier Disease |
| | | | | Sakuntabhai, A. et al. |
| | | | | Mutations in ATP2A2, encoding a Ca2+ pump, cause Darier disease |
| | | | | Nat Genet 21, 271-7 (1999). |
| | | 586225\|Atp2a2 | 1.3E-42 | [Mus musculus][Active transporter, primary; Hydrolase; Transporter; ATPase][Unspecified membrane] Sarcoplasmic reticulum Ca(2+)-ATPase 2, slow twich muscle, cardiac and nonmuscle form, pumps calcium from cytoplasm into ER; associated with dilated cardiomyopathy and gene mutations in human ATP2A2 are associated with Darier Disease |
| | | | | Reed, T. D. et al. |
| | | | | The expression of SR calcium transport ATPase and the Na(+)/Ca(2+)Exchanger are antithetically regulated during mouse cardiac development and in Hypo/hyperthyroidism. |
| | | | | J Mol Cell Cardiol 32, 453-64 (2000). |
| 46 | 7510654CD1 | g7018306 | 8.8E-171 | [Homo sapiens] glucose transporter |
| | | | | Ibberson, M. et al. |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues |
| | | | | J. Biol. Chem. 275, 4607-4612 (2000) |
| | | 569540\|SLC2A8 | 7.6E-172 | [Homo sapiens][Active transporter, secondary; Major Facilitator Superfamily; Transporter][Plasma membrane] Solute carrier family 2 member 8 (glucose transporter X1), glucose transporter that may play a role in glucose sensing |
| | | | | Ibberson, M. et al. (supra) |
| | | | | Proc Natl Acad Sci U S A 97, 7313-8 (2000). |
| | | 757694\|Slc2a8 | 1.9E-150 | [Rattus norvegicus][Transporter] Solute carrier family 2 member 8 (glucose transporter X1), glucose transporter associated with streptozotocin diabetes upon upregulation of mRNA but not protein |
| | | | | Reagan, L. P. et al. |
| | | | | Localization and regulation of GLUTx1 glucose transporter in the hippocampus of streptozotocin diabetic rats. |
| | | | | Proc Natl Acad Sci U S A 98, 2820-5. (2001). |
| 47 | 7510660CD1 | g12248394 | 0.0 | [Mus musculus] cation-transporting atpase |
| | | 610956\|CGI-152 | 0.0 | [Homo sapiens][Active transporter, primary; Hydrolase; Transporter; ATPase] Member of the E1-E2 ATPase family of cation transporters, has a region of weak similarity to a region of rat Atp1a2, which is the catalytic subunit of the sodium- and potassium-transporting ATPase |
| | | 239097\|C10C6.6 | 0.0 | [Caenorhabditis elegans][Active transporter, primary; Hydrolase; Transporter; ATPase][Unspecified membrane] Member of the P-type ATPase, Ca2+-type subfamily protein family |
| 48 | 7510661CD1 | g12248394 | 0.0 | [Mus musculus] cation-transporting atpase |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 610956\|CGI-152 | 5.7E-254 | [Homo sapiens][Active transporter, primary; Hydrolase; Transporter; ATPase] Member of the E1-E2 ATPase family of cation transporters, has a region of weak similarity to a region of rat Atp1a2, which is the catalytic subunit of the sodium- and potassium-transporting ATPase |
| | | 239097\|C10C6.6 | 6.7E-164 | [Caenorhabditis elegans][Active transporter, primary; Hydrolase; Transporter; ATPase][Unspecified membrane] Member of the P-type ATPase, Ca2+-type subfamily protein family |
| 49 | 7510680CD1 | g3901268 | 4.4E-82 | [Rattus norvegicus] SV2 related protein |
| | | | | Janz, R. et al. |
| | | | | SVOP, an evolutionarily conserved synaptic vesicle protein, suggests novel transport functions of synaptic vesicles |
| | | | | J. Neurosci. 18, 9269-9281 (1998) |
| | | 332780\|Rn.30057 | 3.8E-83 | [Rattus norvegicus][Vesicle coat protein; Transporter][Cytoplasmic; Unspecified membrane] Synaptic vesicle protein containing twelve transmembrane domains |
| | | | | Janz, R. et al. (supra) |
| | | 757012\|Slc22a7 | 3.5E-29 | [Rattus norvegicus][Active transporter, secondary; Major Facilitator Superfamily; Transporter][Plasma membrane] Organic cation transporter 2 (solute carrier family 22 member 7), a multispecific sodium-independent organic anion transporter expressed predominantly in the liver, mediates the uptake of salicylate, indomethacin, and nucleoside derivitives |
| | | | | Sekine, T. et al. |
| | | | | Identification of multispecific organic anion transporter 2 expressed predominantly in the liver. |
| | | | | FEBS Lett 429, 179-82 (1998). |
| | | | | Morita, N. et al. |
| | | | | Functional characterization of rat organic anion transporter 2 in LLC-PK1 cells. |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | J Pharmacol Exp Ther 298, 1179-84. (2001). |
| 50 | 7505145CD1 | g15929042 | 2.7E-160 | [Homo sapiens] tetracycline transporter-like protein |
| | | 343916\|TETRAN | 2.3E-161 | [Homo sapiens][Active transporter, secondary; Major Facilitator Superfamily; Transporter][Unspecified membrane] Tetracycline transporter-like protein, member of a superfamily of transporter proteins, may be involved in tetracycline transport |
| | | | | Duyao, M. P. et al. |
| | | | | A gene from chromosome 4p16.3 with similarity to a superfamily of transporter proteins. |
| | | | | Hum Mol Genet 2, 673-6 (1993). |
| 51 | 7505162CD1 | g2765461 | 2.9E-140 | [Homo sapiens] glucose 6-phosphate translocase |
| | | | | Gerin, I. et al. |
| | | | | Sequence of a putative glucose 6-phosphate translocase, mutated in glycogen storage disease type Ib |
| | | | | FEBS Lett. 419, 235-238 (1997) |
| | | 335420\|G6PT1 | 2.5E-141 | [Homo sapiens][Active transporter, secondary; Transporter][Endoplasmic reticulum; Cytoplasmic; Microsomal fraction; Unspecified membrane] Glucose-6-phosphate translocase, component of glucose-6-phosphatase enzyme complex, involved in glycogen metabolism, inhibited by chlorogenic acid and its synthetic derivatives; deficiency is a cause of glycogen storage disease type Ib, Ic, and Id |
| | | | | Kure, S. et al. |
| | | | | Molecular analysis of glycogen storage disease type Ib: identification of a prevalent mutation among Japanese patients and assignment of a putative glucose-6-phosphate translocase gene to chromosome 11. |
| | | | | Biochem Biophys Res Commun 248, 426-31 (1998). |
| | | | | Narisawa, K. et al. |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | A new variant of glycogen storage disease type I probably due to a defect in the glucose-6-phosphate transport system. |
| | | | | Biochem Biophys Res Commun 83, 1360-4 (1978). |
| | | | | Gerin, I. et al. |
| | | | | Sequence of a putative glucose 6-phosphate translocase, mutated in glycogen storage disease type Ib |
| | | | | FEBS Lett 419, 235-8 (1997). |
| | | | | Hou, D. C. et al. |
| | | | | Glycogen storage disease type Ib: structural and mutational analysis of the microsomal glucose-6-phosphate transporter gene. |
| | | | | Am J Med Genet 86, 253-7. (1999). |
| | | 711464\|G6pt1 | 1.1E-133 | [Rattus norvegicus][Active transporter, secondary; Transporter][Endoplasmic reticulum; Cytoplasmic; Unspecified membrane] Glucose-6-phosphate translocase, component of glucose-6-phosphatase enzyme complex, inhibited by chlorogenic acid and its synthetic derivatives, upregulated by insulin deficiency and hyperglycemia in streptozotocin-induced diabetes |
| | | | | Lin, B. et al. |
| | | | | Cloning and characterization of cDNAs encoding a candidate glycogen storage disease type 1b protein in rodents. |
| | | | | J Biol Chem 273, 31656-60 (1998). |
| | | | | Li, Y. et al. |
| | | | | Diabetes affects similarly the catalytic subunit and putative glucose-6-phosphate translocase of glucose-6-phosphatase. |
| | | | | J Biol Chem 274, 33866-8 (1999). |
| 52 | 7505469CD1 | g13111752 | 1.8E-76 | [Homo sapiens] solute carrier family 7 (cationic amino acid transporter, y+ system), member 7 |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 742298\|SLC7A6 | 6.3E-113 | [Homo sapiens] Protein with high similarity to solute carrier family 7 member 7 (human SLC7A7), which is a cationic and dibasic amino acid transporter associated with lysinuric protein intolerance, member of the amino acid permease family of membrane transporters |
| | | 338114\|SLC7A7 | 1.5E-77 | [Homo sapiens][Active transporter, secondary; Transporter][Plasma membrane] Solute carrier family 7 (y+L amino acid transporter-1) member 7, a cationic and dibasic amino acid transporter that forms a heterodimer with the 4F2 heavy chain (SLC3A2); mutation in the corresponding gene causes lysinuric protein intolerance |
| | | | | Mykkanen, J. et al. |
| | | | | Functional analysis of novel mutations in y(+)LAT-1 amino acid transporter gene causing lysinuric protein intolerance (LPI). |
| | | | | Hum Mol Genet 9, 431-8. (2000). |
| 53 | 7505475CD1 | g17223622 | 0.0 | [Homo sapiens] ATP-binding cassette A6 |
| | | 762515\|ABCA9 | 1.5E-189 | [Homo sapiens] Member of the ABC transporter family, which are involved in translocation of a variety of compounds across biological membranes, has low similarity to ATP binding cassette subfamily A member 1 (human ABCA1), which is associated with Tangier disease |
| | | 568162\|ABCA8 | 4.1E-180 | [Homo sapiens][ATP-binding cassette; Active transporter, primary; Hydrolase; Transporter; ATPase][Unspecified membrane; Plasma membrane] ATP-binding cassette subfamily A member 8, a putative transporter |
| | | | | Kaminski, W. E. et al. |
| | | | | ABCA6, a novel a subclass ABC transporter. |
| | | | | Biochem Biophys Res Commun 285, 1295-301. (2001). |
| 54 | 7505568CD1 | g9230651 | 2.5E-40 | [Homo sapiens] facilitative glucose transporter family member GLUT9 |
| | | | | Phay, J. E. et al. |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Cloning and expression analysis of a novel member of the facilitative glucose transporter family, SLC2A9 (GLUT9) |
| | | | | Genomics 66, 217-220 (2000) |
| | | 606512\|SLC2A9 | 2.1E-41 | [Homo sapiens][Transporter][Plasma membrane; Unspecified membrane] Solute carrier family 2 (facilitated glucose transporter) member 9, member of the glucose transporter family, a putative plasma membrane protein that may be involved in the transport of carbohydrates, expressed in the kidney and liver |
| | | | | Phay, J. E. et al. |
| | | | | Cloning and expression analysis of a novel member of the facilitative glucose transporter family, SLC2A9 (GLUT9). |
| | | | | Genomics 66, 217-20 (2000). |
| | | | | Doege, H. et al. |
| | | | | Activity and genomic organization of human glucose transporter 9 (GLUT9), a novel member of the family of sugar-transport facilitators predominantly expressed in brain and leucocytes. |
| | | | | Biochem J 350, 771-776 (2000). |
| 55 | 7506953CD1 | g204204 | 3.2E-160 | [Rattus rattus] GABA-A receptor alpha-2 subunit |
| | | | | Wisden, W. et al. |
| | | | | The distribution of 13 GABAA receptor subunit mRNAs in the rat brain. I. Telencephalon, diencephalon, mesencephalon |
| | | | | J. Neurosci. 12, 1040-1062 (1992) |
| | | | | Laurie, D. J. et al. |
| | | | | The distribution of 13 GABAA receptor subunit mRNAs in the rat brain. II. Olfactory bulb and cerebellum |
| | | | | J. Neurosci. 12, 1063-1076 (1992) |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 339368\|GABRA2 | 9.3E-172 | [Homo sapiens][Channel (passive transporter); Receptor (signaling); Transporter][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Hadingham, K. L. et al. (supra) |
| | | | | Loup, F. et al. (supra) |
| | | 582959\|Gabra2 | 8.6E-167 | [Mus musculus][Channel (passive transporter); Transporter; Receptor (signaling)][Plasma membrane] Alpha 2 subunit of the gamma-aminobutyric acid type A (GABA-A) receptor, a putative chloride channel that is the major inhibitory neurotransmitter receptor in the brain |
| | | | | Sibille, E. et al. (supra) |
| | | | | Bouilleret, V. et al. (supra) |
| 56 | 7510176CD1 | g12620132 | 6.9E-11 | [Homo sapiens] renal sodium/sulfate cotransporter |
| | | | | Lee, A. et al. |
| | | | | The Human Renal Sodium Sulfate Cotransporter (SLC13A1; hNaSi-1) cDNA and Gene: Organization, Chromosomal Localization, and Functional Characterization |
| | | | | Genomics 70, 354-363 (2000) |
| | | 657791\|SLC13A1 | 5.9E-12 | [Homo sapiens] Protein with strong similarity to sodium/sulfate cotransporter 2 (rat Slc13a1), which mediates sodium-dependent transport of sulfate in brush border cells of renal proximal tubules, member of the Sodium:sulfate symporter family of membrane transporters |
| | | | | Lee A et al. |
| | | | | The human renal sodium sulfate cotransporter (SLC13A1; hNaSi-1) cDNA and gene: organization, chromosomal localization, and functional characterization. |
| | | | | Genomics 70, 354-63 (2000). |
| 57 | 7510541CD1 | g10242111 | 8.2E-49 | [Homo sapiens] Na+ and H+ coupled amino acid transport system N |

# Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | Fei, Y. J. et al. |
| | | | | Primary structure, genomic organization, and functional and electrogenic characteristics of human system N 1, a Na+- and H+-coupled glutamine transporter |
| | | | | J. Biol. Chem. 275, 23707-23717 (2000) |
| | | 428246\|SLC38A3 | 7.0E-50 | [Homo sapiens][Transporter][Plasma membrane] Member of the transmembrane amino acid transporter (permease) family |
| | | | | Gu, S. et al. |
| | | | | Identification and characterization of an amino acid transporter expressed differentially in liver. |
| | | | | Proc Natl Acad Sci U S A 97, 3230-5 (2000). |
| | | | | Nakanishi, T. et al. |
| | | | | Structure, function, and tissue expression pattern of human sn2, a subtype of the amino acid transport system n. |
| | | | | Biochem Biophys Res Commun 281, 1343-8. (2001). |
| | | 749242\|SLC38A5 | 6.2E-16 | [Homo sapiens] Solute carrier family 38 member 5 (amino acid transport system N2), a system N amino acid transporter that mediates transport of neutral specific amino acids glutamine, asparagine, and histidine, as well as the transport of serine, alanine, and glycine |
| | | | | Nakanishi, T. et al. (supra) |
| 58 | 7510923CD1 | g1840045 | 3.3E-148 | [Homo sapiens] transporter protein |
| | | 428246\|SLC38A3 | 2.8E-149 | [Homo sapiens][Transporter][Plasma membrane] Member of the transmembrane amino acid transporter (permease) family |
| | | | | Gu, S. et al. (supra) |
| | | | | Nakanishi, T. et al. (supra) |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | 749242\|SLC38A5 | 1.4E-67 | [Homo sapiens] Solute carrier family 38 member 5 (amino acid transport system N2), a system N amino acid transporter that mediates transport of neutral specific amino acids glutamine, asparagine, and histidine, as well as the transport of serine, alanine, and glycine |
| | | | | Nakanishi, T. et al. (supra) |
| 59 | 7510984CD1 | g3643190 | 0.0 | [Homo sapiens] sulfonylurea receptor 1 |
| | | 338358\|ABCC8 | 0.0 | [Homo sapiens][ATP-binding cassette; Hydrolase; Channel (passive transporter); Transporter; Receptor (signaling); ATPase][Plasma membrane] ATP-binding cassette subfamily C member 8, sulfonylurea receptor and subunit of a potassium channel, regulates insulin secretion and potassium transport; gene mutations are implicated in familial persistent hyperinsulinemic hypoglycemia in infancy |
| | | | | Inagaki, N. et al. |
| | | | | Reconstitution of IKATP: an inward rectifier subunit plus the sulfonylurea receptor |
| | | | | Science 270, 1166-70 (1995). |
| | | | | Thomas, P. M. et al. |
| | | | | Mutations in the sulfonylurea receptor gene in familial persistent hyperinsulinemic hypoglycemia of infancy |
| | | | | Science 268, 426-9 (1995). |
| | | 590671\|Abcc8 | 0.0 | [Rattus norvegicus][ATP-binding cassette; Hydrolase; Channel (passive transporter); Transporter; Receptor (signaling); ATPase][Plasma membrane] ATP-binding cassette subfamily C member 8, sulfonylurea receptor and subunit of a potassium channel, regulates potassium transport; mutations in human ABCC8 are linked to familial persistent hyperinsulinemic hypoglycemia in infancy |
| | | | | Aguilar-Bryan, L. et al. |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: or PROTEOME ID NO: | Probability Score | Annotation |
|---|---|---|---|---|
| | | | | |
| | | | | Cloning of the beta cell high-affinity sulfonylurea receptor: a regulator of insulin secretion |
| | | | | Science 268, 423-6 (1995). |
| | | | | Thomas, P. M. et al. (supra) |
| | | | | Malhi, H. et al. |
| | | | | KATP channels regulate mitogenically induced proliferation in primary rat hepatocytes and human liver cell lines. Implications for liver growth control and potential therapeutic targeting. |
| | | | | J Biol Chem 275, 26050-7 (2000). |

# Table 3

<table>
<tr><th>SEQ ID NO:</th><th>Incyte Polypeptide ID</th><th>Amino Acid Residues</th><th>Potential Phosphorylation Sites</th><th>Potential Glycosylation Sites</th><th>Signature Sequences, Domains and Motifs</th><th>Analytical Methods and Databases</th></tr>
<tr><td>1</td><td>7509332CD1</td><td>195</td><td>S26 S48 S126 S146 S156 T70 T78 T101</td><td>N111 N151</td><td>Signal_cleavage: M1-A22</td><td>SPSCAN</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M8-A22</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M7-P24</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M3-Q23</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M1-A22</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M3-A27</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M3-A22</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M3-A29</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Folate receptor family: M7-S195</td><td>HMMER_PFAM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>PROTEIN FOLATE RECEPTOR GLYCOPROTEIN PRECURSOR SIGNAL FOLATEBINDING MEMBRANE GPIANCHOR MULTIGENE PD006906: E50-I192, P24-Y100</td><td>BLAST_PRODOM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>FOLATE-BINDING PROTEIN DM02165|P15328|22-256: D87-P188, A22-C102</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>FOLATE-BINDING PROTEIN DM02165|P02702|1-221: D87-I192, P24-C102</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>FOLATE-BINDING PROTEIN DM02165|P41439|2-242: K41-S195, A4-Y100</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>FOLATE-BINDING PROTEIN DM02165|P14207|2-254: D87-G191, W5-G155</td><td>BLAST_DOMO</td></tr>
<tr><td>2</td><td>7509102CD1</td><td>138</td><td>S54 S81 S117 T56 T136</td><td>N79 N114 N134</td><td>Signal_cleavage: M1-G26</td><td>SPSCAN</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M1-T22</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M1-A24</td><td>HMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Signal Peptide: M1-A29</td><td>HMMER</td></tr>
</table>

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Neurotransmitter-gated ion-channel ligand binding: E59-K135 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: N71-N108, I125-N134 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: T91-W107, M124-K135 | BLIMPS_PRINTS |
| | | | | | Nicotinic acetylcholine receptor signature PR00254: P78-V94, F112-W126, I130-R138 | BLIMPS_PRINTS |
| | | | | | Luteovirus ORF6 protein signature PR00910: P40-G52 | BLIMPS_PRINTS |
| | | | | | NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA2 CHAIN PRECURSOR POSTSYNAPTIC MEMBRANE IONIC CHANNEL GLYCOPROTEIN SIGNAL TRANSMEMBRANE MULTIGENE FAMILY PD108915: M1-T58 | BLAST_PRODOM |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: E57-N134 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|A40110\|16-509: L16-A24, P46-N134 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P09480\|14-526: H55-N134 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P43681\|17-626: S54-N134 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|JC4021\|17-627: S54-N134 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| 3 | 7509132CD1 | 355 | S2 S23 S24 S72 S89 S158 S334 T94 T129 T344 Y348 | N273 | Sugar (and other) transporter: C5-F329 | HMMER_PFAM |
| | | | | | Cytosolic domains: P70-P133, E191-R196, I252-P263, R313-D355; Transmembrane domains: L47-C69, L134-F153, P168-V190, T197-L219, V229-F251, A264-Y286, A290-L312; Non-cytosolic domains: M1-S46, Y154-Q167, L220-Y228, V287-E289 | TMHMMER |
| | | | | | Sugar transport proteins signatures: Y170-A225 | PROFILESCAN |
| | | | | | Sugar transporter signature PR00171: Q144-Y154, I231-I252, A254-M266 | BLIMPS_PRINTS |
| | | | | | Glucose transporter signature PR00172: L134-Y155, A169-V190, L200-L220, I231-A254, A264-M282, Y294-V314 | BLIMPS_PRINTS |
| | | | | | GLUCOSE TRANSPORTER TYPE INSULIN RESPONSIVE DUPLICATION TRANSMEMBRANE SUGAR TRANSPORT GLYCOPROTEIN MULTIGENE PD015687: G319-D355 | BLAST_PRODOM |
| | | | | | GLUCOSE TRANSPORTER TYPE 3 CEFGT3 DUPLICATION TRANSMEMBRANE SUGAR TRANSPORT GLYCOPROTEIN MULTIGENE FAMILY PD073462: A288-E353 | BLAST_PRODOM |
| | | | | | SUGAR TRANSPORT PROTEINS DM00135\|P14672\|122-474: Q34-T321 | BLAST_DOMO |
| | | | | | SUGAR TRANSPORT PROTEINS DM00135\|P19357\|122-474: Q34-T321 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | SUGAR TRANSPORT PROTEINS DM00135\|P14142\|124-475: Q34-T321 | BLAST_DOMO |
| | | | | | SUGAR TRANSPORT PROTEINS DM00135\|P27674\|126-458: Q34-T321 | BLAST_DOMO |
| | | | | | Sugar transport proteins signature 1: S186-G202 | MOTIFS |
| 4 | 7509136CD1 | 380 | S2 S28 S66 S291 T129 | N92 | Neurotransmitter-gated ion-channel ligand binding domain: M1-P173 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channel transmembrane region: V180-F366 | HMMER_PFAM |
| | | | | | Cytosolic domain: P198-R348; Transmembrane domains: F175-L197, L349-Y371; Non-cytosolic domains: M1-L174, H372-P380 | TMHMMER |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: V36-N45, D64-Y102, R160-A201 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion-channels signature: V59-Q110 | PROFILESCAN |
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: S2-W18, S35-N46, C79-C93, L167-N179 | BLIMPS_PRINTS |
| | | | | | Nicotinic acetylcholine receptor signature PR00254: H23-W37, V41-V53, V59-S77 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: S2-A201, P302-F366, V178-D247 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P02712\|7-486: M1-A201, I185-P374 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P04758\|7-498: M1-Q287, I185-P374 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P04759\|4-495: M1-D200, Q161-P374 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P02718\|4-501: M1-A201, L186-P374 | BLAST_DOMO |
| | | | | | Leucine zipper pattern: L190-L211 | MOTIFS |
| | | | | | Neurotransmitter-gated ion-channels signature: C79-C93 | MOTIFS |
| 5 | 7509178CD1 | 375 | S21 S81 S271 S290 S312 S326 T266 T323 Y319 | N79 | Signal_cleavage: M1-G20 | SPSCAN |
| | | | | | Signal Peptide: M1-A15, M1-G20, M1-H23 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: Q78-P149, E24-K77 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channel transmembrane region: V156-F364 | HMMER_PFAM |
| | | | | | Cytosolic domains: P174-S184, N235-H346; Transmembrane domains: Y151-L173, I185-I202, I212-I234, I347-I369; Non-cytosolic domains: M1-L150, P203-L211, E370-G375 | TMHMMER |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236:V51-Y89, Y136-S177 | BLIMPS_BLOCKS |
| | | | | | Nicotinic acetylcholine receptor signature PR00254:I58-V74 | BLIMPS_PRINTS |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Metabotropic glutamate receptor signature PR00593: T182-L196 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: K77-K338, S287-F364, E22-W87 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P25108\|4-453: Q68-N372, L6-I116 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P22456\|4-453: Q78-N372, L6-R120 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P02711\|8-457: Q68-N372, L6-I116 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|S60589\|5-492: Q78-S271, P308-F364, L7-S92 | BLAST_DOMO |
| 6 | 7509214CD1 | 153 | S128 T138 | | Signal Peptide: M1-G18 | HMMER |
| | | | | | PERIPHERAL BENZODIAZEPINE RECEPTOR RELATED PROTEIN PD068564: M52-S153 | BLAST_PRODOM |
| 7 | 7509244CD1 | 369 | S2 S28 S66 T129 | N92 | Neurotransmitter-gated ion-channel ligand binding domain: M1-P173 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channel transmembrane region: V180-F355 | HMMER_PFAM |
| | | | | | Cation transporter family protein: M1-A358 | HMMER_TIGRFAM |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Cytosolic domains: P198-M205, H261-R337; Transmembrane domains: F175-L197, G206-A223, K238-L260, L338-Y360; Non-cytosolic domains: M1-L174, D224-I237, H361-P369 | TMHMMER |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236:V36-N45, D64-Y102, R160-A201 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion-channels signature: V59-Q110 | PROFILESCAN |
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: S2-W18, S35-N46, C79-C93, L167-N179 | BLIMPS_PRINTS |
| | | | | | Gamma-aminobutyric acid (GABA) receptor signature PR00253: Y176-Y196, A201-L222, L231-I252 | BLIMPS_PRINTS |
| | | | | | Nicotinic acetylcholine receptor signature PR00254: H23-W37, V41-V53, V59-S77 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: S2-S309, P291-F355 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P04758\|7-498: M1-E300, L297-P363 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P02712\|7-486: M1-L301, L297-P363 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P09484\|2-474: M1-K292, P291-Y360 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P02713\|5-498: M1-R313, P299-H361 | BLAST_DOMO |
| | | | | | Neurotransmitter-gated ion-channels signature: C79-C93 | MOTIFS |
| 8 | 7509256CD1 | 303 | S26 S27 S44 S93 S174 T97 T194 T207 T254 T261 Y206 Y240 | N54 N242 | Signal_cleavage: M1-S22 | SPSCAN |
| | | | | | Signal Peptide: M1-L15 | HMMER |
| | | | | | Signal Peptide: M1-E19 | HMMER |
| | | | | | Signal Peptide: M1-A20 | HMMER |
| | | | | | Signal Peptide: M1-S22 | HMMER |
| | | | | | Signal Peptide: M1-W16 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: T56-V266 | HMMER_PFAM |
| | | | | | Cation transporter family protein: L5-W303 | HMMER_TIGRFAM |
| | | | | | Cytosolic domain: I290-W303; Transmembrane domain: G267-W289; Non-cytosolic domain: M1-V266 | TMHMMER |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: V82-P119, L138-N147, D168-Y206, Y253-A294 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion-channels signature: L163-S217 | PROFILESCAN |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: R102-L118, C137-E148, C183-C197, F260-G272 | BLIMPS_PRINTS |
| | | | | | Gamma-aminobutyric acid (GABA) receptor signature PR00253: Y269-W289 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: I59-G302 | BLAST_PRODOM |
| | | | | | GLYCINE RECEPTOR BETA CHAIN PRECURSOR POSTSYNAPTIC MEMBRANE IONIC CHANNEL GLYCOPROTEIN PD022977: M1-N58 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P48167\|13-497: I13-G302 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|S18836\|18-453: E19-G302 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|B49970\|18-452: E19-D131, G30-G302 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P23415\|12-449: L38-G302 | BLAST_DOMO |
| | | | | | Neurotransmitter-gated ion-channels signature: C183-C197 | MOTIFS |
| 9 | 7509395CD1 | 370 | S21 S76 S266 S285 S307 S321 T261 T318 Y314 | N74 | Signal_cleavage: M1-G20 | SPSCAN |
| | | | | | Signal Peptide: M1-A15 | HMMER |
| | | | | | Signal Peptide: M1-G20 | HMMER |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Signal Peptide: M1-H23 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: N74-P144, E24-N73 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channel transmembrane region: V151-F359 | HMMER_PFAM |
| | | | | | Cytosolic domains: P169-S179, N230-H341; Transmembrane domains: Y146-L168, I180-I197, I207-I229, I342-I364; Non-cytosolic domains: M1-L145, P198-L206, E365-G370 | TMHMMER |
| | | | | | Cation transporter family protein M1-V358 | HMMER_TIGRFAM |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: R46-Y84, Y131-S172 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: F138-N150 | BLIMPS_PRINTS |
| | | | | | Nicotinic acetylcholine receptor signature PR00254: I58-N74 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: N74-K333, S282-F359, E22-N73 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P25108\|4-453: L6-N73, N74-N367 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P02711\|8-457: L6-N73, N74-N367 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P22456\|4-453: L6-N73, N74-N367 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P18845\|7-510: L6-G289, P303-I364 | BLAST_DOMO |
| 10 | 7503287CD1 | 283 | S40 S104 S146 S180 S199 T62 T75 T209 T239 | N276 | Signal_cleavage: M1-C61 | SPSCAN |
| | | | | | Amiloride-sensitive sodium channel: F21-N283 | HMMER_PFAM |
| | | | | | ENaC: sodium channel transporter: S17-N283 | HMMER_TIGRFAM |
| | | | | | Amiloride-sensitive sodium channels proteins BL01206: A20-L30, Y191-R204, G213-P231 | BLIMPS_BLOCKS |
| | | | | | Amiloride-sensitive sodium channel alpha-subunit signature PR01078: R43-V60, S83-R99, L169-S180, E182-N198, G213-P231 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE ION TRANSPORT SODIUM GLYCOPROTEIN AMILORIDESENSITIVE SUBUNIT NA+ PD001186: F21-P151, R160-I249 | BLAST_PRODOM |
| | | | | | CHANNEL IONIC TRANSMEMBRANE SODIUM ION TRANSPORT PROTON GATED CATION ASIC1 PD151848: N119-D159 | BLAST_PRODOM |
| | | | | | SODIUM; SENSITIVE; AMILORIDE; CHANNEL; DM01114\|P51169\|10-589: F21-E123, G162-P231 | BLAST_DOMO |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | SODIUM; SENSITIVE; AMILORIDE; CHANNEL; DM01114\|P51167\|1-556: F21-L108, G162-V232 | BLAST_DOMO |
| | | | | | SODIUM; SENSITIVE; AMILORIDE; CHANNEL; DM01114\|P51170\|7-576: F21-Q134, H163-P231 | BLAST_DOMO |
| | | | | | SODIUM; SENSITIVE; AMILORIDE; CHANNEL; DM01114\|P51171\|8-580: I18-I137, G162-P231 | BLAST_DOMO |
| 11 | 7503320CD1 | 90 | S84 | N46 | Signal_cleavage: M1-Q22 | SPSCAN |
| | | | | | Signal Peptide: M1-V19 | HMMER |
| | | | | | Signal Peptide: M1-G23 | HMMER |
| | | | | | Signal Peptide: M1-F25 | HMMER |
| | | | | | Signal Peptide: M1-Q22 | HMMER |
| | | | | | Nicotinic acetylcholine receptor signature PR00254: L60-T76 | BLIMPS_PRINTS |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P54131\|3-491: G6-T74 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|JH0173\|14-503: V8-T74 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P48180\|2-497: L10-L72 | BLAST_DOMO |
| 12 | 7503335CD1 | 549 | S6 S47 S86 S131 S344 S356 S374 S444 T15 T28 T124 T149 T204 T351 T421 T462 T509 | N187 N202 N213 | ATP P2X receptor: F11-V358 | HMMER_PFAM |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | P2X: cation transporter protein: M1-C328 | HMMER_TIGRFAM |
| | | | | | Cytosolic domain: S47-Q286; Transmembrane domains: M24-V46, L287-I309; Non-cytosolic domains: M1-S23, D310-Y549 | TMHMMER |
| | | | | | ATP P2X receptors proteins BL01212: F38-T90, C129-V153, T161-E171, R261-V288 | BLIMPS_BLOCKS |
| | | | | | P2X PURINOCEPTOR ATP RECEPTOR P2X7 PURINERGIC P2Z IONIC CHANNEL TRANSMEMBRANE PD041643: I309-Y549 | BLAST_PRODOM |
| | | | | | RECEPTOR ATP IONIC CHANNEL TRANSMEMBRANE ION TRANSPORT P2X PURINOCEPTOR PURINERGIC PD002383: F11-Q243, E335-K349, G211-I309 | BLAST_PRODOM |
| | | | | | PORE-FORMING MOTIF DOMAIN DM02085\|P51577\|1-384: M1-Q243, W245-I309, C331-E347 | BLAST_DOMO |
| | | | | | PORE-FORMING MOTIF DOMAIN DM02085\|P51579\|1-378: Y13-I234, W245-I309 | BLAST_DOMO |
| | | | | | PORE-FORMING MOTIF DOMAIN DM02085\|P51578\|1-385: V10-Q243, W245-I309 | BLAST_DOMO |
| | | | | | PORE-FORMING MOTIF DOMAIN DM02085\|P51575\|1-380: V10-E237, G200-I309 | BLAST_DOMO |
| 13 | 7503952CD1 | 246 | S98 S190 S205 T172 T179 Y134 | N52 N96 N138 N168 N203 N232 | Signal_cleavage: M1-A21 | SPSCAN |
| | | | | | Signal Peptide: M6-A21 | HMMER |
| | | | | | Signal Peptide: M6-D23 | HMMER |
| | | | | | Signal Peptide: M1-A21 | HMMER |
| | | | | | Signal Peptide: M1-T24 | HMMER |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Signal Peptide: M1-P27 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: L32-P237 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: V59-N96, I113-E122, S140-H178 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion-channels signature: V135-R189 | PROFILESCAN |
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: K79-W95, A112-F123, C155-C169 | BLIMPS_PRINTS |
| | | | | | Nicotinic acetylcholine receptor signature PR00254: H66-V82, F100-W114, V135-S153 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: L35-N232 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P46098\|7-476: A31-F231 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P12391\|15-459: W10-S211 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P54131\|3-491: C12-V212 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|A40110\|16-509: L9-S217 | BLAST_DOMO |
| 14 | 7504530CD1 | 273 | S10 T11 T190 | | Major intrinsic protein: E27-H251 | HMMER_PFAM |
| | | | | | MIP family channel proteins: A39-V273 | HMMER_TIGRFAM |

Table 3

<table>
<tr><th>SEQ ID NO:</th><th>Incyte Polypeptide ID</th><th>Amino Acid Residues</th><th>Potential Phosphorylation Sites</th><th>Potential Glycosylation Sites</th><th>Signature Sequences, Domains and Motifs</th><th>Analytical Methods and Databases</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Cytosolic domains: M1-E36, H92-K111, L162-T167, N226-V273; Transmembrane domains: F37-V59, G69-A91, F112-F134, L139-Y161, L168-T190, A203-I225; Non-cytosolic domains: L60-L68, Y135-I138, D191-E202</td><td>TMHMMER</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>MIP family proteins BL00221: A39-V49, I88-T98, E175-D191, T221-I235, T237-F247</td><td>BLIMPS_BLOCKS</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Major intrinsic protein family signature PR00783: R35-S54, F74-T98, K111-I130, N174-Q192, G207-R229</td><td>BLIMPS_PRINTS</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>TRANSMEMBRANE TRANSPORT PROTEIN AQUAPORIN INTRINSIC CHANNEL MEMBRANE WATER TONOPLAST FAMILY PD000295: R35-H259</td><td>BLAST_PRODOM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>AQUAPORIN7 LIKE AQUAPORIN ADIPOSE AQPAP TRANSPORT TRANSMEMBRANE PD062309: M1-V34</td><td>BLAST_PRODOM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>MIP FAMILY DM00228|P47862|15-263: L29-V273</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>MIP FAMILY DM00228|I59266|15-263: L29-V273</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>MIP FAMILY DM00228|P43549|340-587: R31-G272</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>MIP FAMILY DM00228|P11244|1-253: L38-G272</td><td>BLAST_DOMO</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Prenyl group binding site (CAAX box):</td><td>MOTIFS</td></tr>
<tr><td>15</td><td>7509303CD1</td><td>245</td><td>S54 S122 S176 S209 T42 T157</td><td>N64 N65</td><td>Amiloride-sensitive sodium channel: F62-L228</td><td>HMMER_PFAM</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Cytosolic domain: M1-T84; Transmembrane domain: A85-G107; Non-cytosolic domain: E108-Q245</td><td>TMHMMER</td></tr>
</table>

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Amiloride-sensitive sodium channels proteins BL01206: F61-A71 | BLIMPS_BLOCKS |
| | | | | | Amiloride-sensitive sodium channel alpha-subunit signature PR01078: T84-Q101, D123-R139 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE ION TRANSPORT SODIUM GLYCOPROTEIN AMILORIDESENSITIVE SUBUNIT NA+ PD001186: F62-V217 | BLAST_PRODOM |
| | | | | | AMILORIDESENSITIVE SODIUM CHANNEL ALPHASUBUNIT NA+ ALPHA SUBUNIT ENAC NONVOLTAGEGATED SCNEA PD040285: M1-F61 | BLAST_PRODOM |
| | | | | | SODIUM; SENSITIVE; AMILORIDE; CHANNEL; DM01114\|A49585\|37-598: A37-L228 | BLAST_DOMO |
| | | | | | SODIUM; SENSITIVE; AMILORIDE; CHANNEL; DM01114\|P37088\|37-598: A37-L228 | BLAST_DOMO |
| | | | | | SODIUM; SENSITIVE; AMILORIDE; CHANNEL; DM01114\|P55270\|17-579: P38-L228 | BLAST_DOMO |
| | | | | | SODIUM; SENSITIVE; AMILORIDE; CHANNEL; DM01114\|P37089\|63-626: A37-L228 | BLAST_DOMO |
| 16 | 7509910CD1 | 364 | S6 S47 S86 S131 S274 T15 T28 T124 T149 T204 | N187 N202 N213 N241 N284 | Signal_cleavage: M1-A44 | SPSCAN |
| | | | | | ATP P2X receptor: F11-V347 | HMMER_PFAM |
| | | | | | Cation transporter protein: M1-W358 | HMMER_TIGRFAM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Cytosolic domain: M1-M24; Transmembrane domain: N25-S47; Non-cytosolic domain: D48-D364 | TMHMMER |
| | | | | | ATP P2X receptors proteins BL01212: F38-T90, C129-V153, T161-E171, A185-I208, P224-L278, P289-Y299, R307-V334 | BLIMPS_BLOCKS |
| | | | | | RECEPTOR ATP IONIC CHANNEL TRANSMEMBRANE ION TRANSPORT P2X PURINOCEPTOR PURINERGIC PD002383: F11-L346 | BLAST_PRODOM |
| | | | | | PORE-FORMING MOTIF DOMAIN DM02085\|P51577\|1-384: M1-L346 | BLAST_DOMO |
| | | | | | PORE-FORMING MOTIF DOMAIN DM02085\|P51575\|1-380: V10-L346 | BLAST_DOMO |
| | | | | | PORE-FORMING MOTIF DOMAIN DM02085\|P49654\|1-370: S6-V335 | BLAST_DOMO |
| | | | | | PORE-FORMING MOTIF DOMAIN DM02085\|P51578\|1-385: V10-G345 | BLAST_DOMO |
| | | | | | ATP P2X receptors signature: G249-F275 | MOTIFS |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| 17 | 7509982CD1 | 1623 | S30 S50 S134 S249 S353 S491 S672 S761 S792 S815 S825 S921 S929 S960 S1041 S1133 S1199 S1275 S1301 S1335 S1494 T111 T206 T558 T572 T624 T643 T755 T772 T780 T858 T974 T1178 T1263 T1346 T1376 T1424 T1447 T1468 T1551 T1611 Y953 | N71 N84 N91 N109 N130 N241 N436 N544 N576 N917 N946 N996 N1311 | Signal peptide: M26-A45, M26-M51 | HMMER |
| | | | | | ABC transporter: G507-G689, G1319-G1495 | HMMER_PFAM |
| | | | | | ATPases associated with a variety of cellular activities: E1318-R1496, E506-R690 | HMMER_INCY |
| | | | | | 3a0106s01: sulfate transport system permease protein: I478-I717, C1299-V1526 | HMMER_TIGRFAM |
| | | | | | 3a0501s02: Type II (General) Secretory Pathway (IISP) Family protein: A477-L691 | HMMER_TIGRFAM |
| | | | | | cbiO: cobalt transport protein ATP-bin: G492-A676 | HMMER_TIGRFAM |
| | | | | | drrA: daunorubicin resistance ABC transporter ATP-binding subunit: K1302-K1597, K485-G790 | HMMER_TIGRFAM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | nodI: nodulation ABC transporter NodI: G474-E795, S1286-K1590 | HMMER_TIGRFAM |
| | | | | | ntrCD: nitrate transport ATP-binding subunits C and D: A1309-F1537, L497-L710 | HMMER_TIGRFAM |
| | | | | | thiQ: ABC transporter, ATP-binding protein, ThiQ subfamily: I478-S697 | HMMER_TIGRFAM |
| | | | | | Cytosolic domains: M1-E29, K244-F263, S319-K324, D418-K859, D1044-A1062, S1123-S1133, R1183-D1202; Transmembrane domains: S30-S49, E221-T243, W264-I286, G296-L318, A325-F347, T395-F417, V860-Y882, F1024-S1043, Y1063-I1085, I1100-I1122, G1134-F1156, I1160-V1182, F1203-L1225; Non-cytosolic domains: S50-N220, T287-T295, Y348-Y394, A883-S1023, F1086-Q1099, D1157-S1159, K1226-P1623 | TMHMMER |
| | | | | | ABC Transporters family signature: V595-D646 | PROFILESCAN |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P41233\|839-1045: I478-N688, K1306-M1492 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P34358\|611-816: I478-N688, A1308-M1492 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P41233\|1851-2058: K1302-S1494, I478-N688 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P23703\|41-246: L500-G689, E1291-G1495 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | ATP/GTP-binding site motif A (P-loop): G514-S521, G1326-S1333 | MOTIFS |
| 18 | 7510082CD1 | 611 | S174 S281 S302 S572 T195 T369 T594 | N61 N66 N178 N223 N356 N439 N597 | POT family: Y101-S259, R362-S503 | HMMER_PFAM |
| | | | | | Cytosolic domains: M1-A37, D95-R100, N178-R197, T253-Q311, D391-K410, S485-M496, R562-C611; Transmembrane domains: V38-L60, A75-A94, Y101-F123, P155-S177, F198-I220, Y230-I252, V312-F331, Y368-K390, M411-L428, I462-Y484, G497-L519, D539-G561; Non-cytosolic domains: N61-R74, P124-S154, Q221-G229, Q332-S367, E429-Q461, P520-M538 | TMHMMER |
| | | | | | PTR2 family proton/oligopeptide symporters proteins BL01022: E42-L60, A72-L117, G164-V187, F199-V211, E472-S507 | BLIMPS_BLOCKS |
| | | | | | TRANSPORTER TRANSPORT TRANSMEMBRANE PEPTIDE OLIGOPEPTIDE PROTEIN SYMPORT ISOFORM H+/PEPTIDE COTRANSPORTER PD001550: Y101-S507 | BLAST_PRODOM |
| | | | | | PEPTIDE/HISTIDINE TRANSPORTER PD127516: S503-R565 | BLAST_PRODOM |
| | | | | | PTR2 FAMILY PROTON/OLIGOPEPTIDE SYMPORTERS DM01990\|P46032\|46-551: E42-S518 | BLAST_DOMO |
| | | | | | PTR2 FAMILY PROTON/OLIGOPEPTIDE SYMPORTERS DM01990\|Q05085\|32-539: A33-N271, S367-H525, D306-V442 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | PTR2 FAMILY PROTON/OLIGOPEPTIDE SYMPORTERS DM01990\|P46031\|84-554: V38-I349, V381-L509 | BLAST_DOMO |
| | | | | | PTR2 FAMILY PROTON/OLIGOPEPTIDE SYMPORTERS DM01990\|P32901\|81-545: V38-V261, A163-F350, I370-L513 | BLAST_DOMO |
| 19 | 7510367CD1 | 55 | | | Mitochondrial energy transfer proteins signature: N25-T53 | PROFILESCAN |
| | | | | | Mitochondrial energy transfer proteins signature: P42-A51 | MOTIFS |
| 20 | 7510413CD1 | 287 | S26 S48 S165 S218 S238 S248 T193 | N163 N203 N243 | Signal_cleavage: M1-A22 | SPSCAN |
| | | | | | Signal Peptide: M8-A22 | HMMER |
| | | | | | Signal Peptide: M7-P24 | HMMER |
| | | | | | Signal Peptide: M3-Q23 | HMMER |
| | | | | | Signal Peptide: M1-A22 | HMMER |
| | | | | | Signal Peptide: M3-A27 | HMMER |
| | | | | | Signal Peptide: M3-A22 | HMMER |
| | | | | | Signal Peptide: M3-A29 | HMMER |
| | | | | | Folate receptor family: M7-S287 | HMMER_PFAM |
| | | | | | PROTEIN FOLATE RECEPTOR GLYCOPROTEIN PRECURSOR SIGNAL FOLATEBINDING MEMBRANE GPIANCHOR MULTIGENE PD006906: P24-Q56, E150-I284 | BLAST_PRODOM |
| | | | | | FOLATE-BINDING PROTEIN DM02165\|P41439\|2-242: A4-Q56, E150-S287 | BLAST_DOMO |
| | | | | | FOLATE-BINDING PROTEIN DM02165\|P14207\|2-254: W5-Q56, T146-G283 | BLAST_DOMO |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | FOLATE-BINDING PROTEIN DM02165\|P15328\|22-256: A22-Q56, T146-P280 | BLAST_DOMO |
| | | | | | FOLATE-BINDING PROTEIN DM02165\|P02702\|1-221: P24-Q56, T146-I284 | BLAST_DOMO |
| 21 | 1721303CD1 | 55 | S3 | | F ATP SYNTHASE CHAIN MITOCHONDRIAL SYNTHESIS HYDROGEN ION TRANSPORT CF0 PD015221: V11-Y54 | BLAST_PRODOM |
| 22 | 7502007CD1 | 272 | S9 T189 | | Major intrinsic protein: E26-H250 | HMMER_PFAM |
| | | | | | MIP: MIP family channel proteins: A38-V272 | HMMER_TIGRFAM |
| | | | | | Cytosolic domains: M1-E35, H91-K110, L161-T166, N225-V272; Transmembrane domains: F36-V58, G68-A90, F111-F133, L138-Y160, L167-T189, A202-I224; Non-cytosolic domains: L59-L67, Y134-I137, D190-E201 | TMHMMER |
| | | | | | MIP family proteins BL00221: A38-V48, I87-T97, E174-D190, T220-I234, T236-F246 | BLIMPS_BLOCKS |
| | | | | | Major intrinsic protein family signature PR00783: R34-S53, F73-T97, K110-I129, N173-Q191, G206-R228 | BLIMPS_PRINTS |
| | | | | | TRANSMEMBRANE TRANSPORT PROTEIN AQUAPORIN INTRINSIC CHANNEL MEMBRANE WATER TONOPLAST FAMILY PD000295: R34-H258 | BLAST_PRODOM |
| | | | | | MIP FAMILY DM00228\|P47862\|15-263: L28-V272 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|I59266\|15-263: L28-V272 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|P43549\|340-587: R30-G271 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|P11244\|1-253: L37-G271 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Prenyl group binding site (CAAX box): | MOTIFS |
| 23 | 7506439CD1 | 188 | T35 T65 | N34 N110 | Signal_cleavage: M1-G27 | SPSCAN |
| | | | | | Signal Peptide: M7-A25 | HMMER |
| | | | | | Signal Peptide: M7-G27 | HMMER |
| | | | | | Signal Peptide: M7-A29 | HMMER |
| | | | | | Signal Peptide: M1-A29 | HMMER |
| | | | | | Signal Peptide: M1-G27 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: L40-R146 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: V67-N104, I121-E130 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: T87-W103, S120-F131 | BLIMPS_PRINTS |
| | | | | | Nicotinic acetylcholine receptor signature PR00254: Y74-I90, F108-W122 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: R42-D125 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P46098\|7-476: Q13-Q148 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P09480\|14-526: R42-F151 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P45963\|9-466: L47-Y142, Y142-P168 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00195\|P26152\|13-440: R42-D133 | BLAST_DOMO |
| | | | | | Leucine zipper pattern: L2-L23 | MOTIFS |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| 24 | 7509243CD1 | 111 | S91 T59 | N57 N86 | Signal_cleavage: M1-A41 | SPSCAN |
| | | | | | Cytosolic domain: Q33-H111; Transmembrane domain: W10-F32; Non-cytosolic domain: M1-C9 | TMHMMER |
| | | | | | TWIK-1 K+ channel subunit signature PR01096: V22-A41, S43-Q63, W64-G80 | BLIMPS_PRINTS |
| 25 | 7509404CD1 | 46 | S26 S27 S41 | | Signal_cleavage: M1-S22 | SPSCAN |
| | | | | | Signal Peptide: M1-L15 | HMMER |
| | | | | | Signal Peptide: M1-E19 | HMMER |
| | | | | | Signal Peptide: M1-A20 | HMMER |
| | | | | | Signal Peptide: M1-S22 | HMMER |
| | | | | | Signal Peptide: M1-W16 | HMMER |
| | | | | | GLYCINE RECEPTOR BETA CHAIN PRECURSOR POSTSYNAPTIC MEMBRANE IONIC CHANNEL GLYCOPROTEIN PD022977: M1-R43 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P48167\|13-497: I13-R43 | BLAST_DOMO |
| 26 | 7509439CD1 | 204 | S37 S47 S98 S184 T27 T32 | N158 | ATP synthase: A26-L141 | HMMER_PFAM |
| | | | | | ATPsyn_F1gamma: ATP synthase F1,: M25-Y204 | HMMER_TIGRFAM |
| | | | | | ATP synthase gamma subunit proteins BL00153: K29-A53, G104-S143 | BLIMPS_BLOCKS |
| | | | | | ATP synthase gamma subunit signature PR00126: K88-H107 | BLIMPS_PRINTS |
| | | | | | GAMMA ATP SYNTHASE CHAIN HYDROLASE SYNTHESIS CF1 HYDROGEN ION TRANSPORT PD001150: A26-R138 | BLAST_PRODOM |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | ATP SYNTHASE GAMMA SUBUNIT DM00493\|P05631\|25-297: M25-Y142 | BLAST_DOMO |
| | | | | | ATP SYNTHASE GAMMA SUBUNIT DM00493\|P49377\|17-288: A26-L141 | BLAST_DOMO |
| | | | | | ATP SYNTHASE GAMMA SUBUNIT DM00493\|P05436\|1-285: M25-R138 | BLAST_DOMO |
| | | | | | ATP SYNTHASE GAMMA SUBUNIT DM00493\|P07227\|1-298: M25-R138 | BLAST_DOMO |
| 27 | 7510202CD1 | 1400 | S2 S7 S61 S70 S113 S197 S625 S755 S778 S788 S884 S892 S923 S1004 S1053 S1096 S1157 S1201 S1261 T30 T86 T458 T463 T575 T716 T743 T756 T792 T937 T966 T970 T1266 T1295 T1338 T1380 Y916 | N72 N120 N195 N244 N456 N545 N556 N880 N909 N959 N1271 N1336 | Signal peptides: M26-S49, M26-D56 | HMMER |
| | | | | | ATPases associated with a variety of cellular activities: E509-K653, K1278-P1400 | HMMER_SMART |
| | | | | | ABC transporter: G510-G652 | HMMER_PFAM |
| | | | | | drrA: daunorubicin resistance ABC transport: K488-G753 | HMMER_TIGRFAM |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Cytosolic domains: M1-T30, N244-A265, K326-F329, H379-Y397, E846-N985, M1045-L1064, D1115-G1120, E1184-P1400; Transmembrane domains: L31-Q53, V221-V243, F266-I288, V303-I325, L330-T352, A356-I378, L398-F420, S823-Y845, T986-I1005, A1025-L1044, L1065-F1087, G1097-T1114, F1121-F1143, I1161-L1183; Non-cytosolic domains: V54-G220, V289-M302, H353-P355, D421-K822, G1006-S1024, I1088-S1096, S1144-E1160 | TMHMMER |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P41233\|839-1045: I481-P610, I1267-L1393, V597-N651 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P34358\|611-816: I481-D602, A1268-V1389, E595-I649 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P36879\|5-211: I481-E578 | BLAST_DOMO |
| | | | | | DM06389\|P41233\|1047-1849: D682-D758, G652-L675, A994-R1090, I796-Y861 | BLAST_DOMO |
| | | | | | ATP/GTP-binding site motif A (P-loop): G517-T524, G1286-S1293 | MOTIFS |
| 28 | 7510203CD1 | 438 | S24 S194 S378 T127 T229 T312 | N23 N65 | ABC transporter transmembrane region: W10-L304 | HMMER_PFAM |
| | | | | | Cytosolic domains: M1-F72, R133-D143, Q188-Q255, F311-G438; Transmembrane domains: Y73-F95, L110-N132, S144-L166, L170-V187, W256-V278, G288-S310; Non-cytosolic domains: A96-R109, G167-G169, Q279-P287 | TMHMMER |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | RESISTANCE; MULTIDRUG; SAUROLEISHMANIA; C3F10.11C; DM01525\|P39109\|861-1270: G50-G357 | BLAST_DOMO |
| | | | | | RESISTANCE; MULTIDRUG; SAUROLEISHMANIA; C3F10.11C; DM01525\|P33527\|868-1291: A7-E27, N65-T328 | BLAST_DOMO |
| | | | | | RESISTANCE; MULTIDRUG; SAUROLEISHMANIA; C3F10.11C; DM01525\|S64757\|872-1300: A7-E27, N65-G357 | BLAST_DOMO |
| | | | | | RESISTANCE; MULTIDRUG; SAUROLEISHMANIA; C3F10.11C; DM01525\|Q10185\|827-1237: F72-S359 | BLAST_DOMO |
| | | | | | Leucine zipper pattern: L145-L166, L152-L173, L159-L180 | MOTIFS |
| 29 | 7510208CD1 | 871 | S21 S50 S119 S140 S199 S256 S281 S467 S502 S533 S631 S823 T16 T48 T252 T353 T382 T440 T612 T633 T696 T799 | N14 N90 N169 N174 N306 N369 N380 N421 N433 N477 N485 N495 N531 N545 N591 N601 N629 | Signal_cleavage: M1-R43 | SPSCAN |
| | | | | | Cytosolic domain: R43-E53; Transmembrane domains: R20-L42, V54-P76; Non-cytosolic domains: M1-R19, D77-A871 | TMHMMER |
| | | | | | ATPBINDING TRANSPORTER CASSETTE ABC GLYCOPROTEIN TRANSMEMBRANE TRANSPORT RIM ABCR SIMILARITY PD006867: Y643-Q664, I663-S744 | BLAST_PRODOM |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Leucine zipper pattern: L509-L530, L516-L537, L534-L555 | MOTIFS |
| | | | | | Cell attachment sequence: R863-D865 | MOTIFS |
| | | | | | Eukaryotic molybdopterin oxidoreductases signature: A391-E425 | MOTIFS |
| 30 | 7510446CD1 | 104 | S70 S86 S90 | | PROTEIN CHLORIDE CHANNEL SKELETAL MUSCLE CLC1 IONIC ION TRANSPORT VOLTAGEGATED PD035113: M1-H57 | BLAST_PRODOM |
| 31 | 7505294CD1 | 336 | S25 S239 S291 T53 T183 | N23 N32 | signal_cleavage: M1-A48 | SPSCAN |
| | | | | | Cytosolic domains: D123-R128, R176-R187, E238-P336<br>Transmembrane domains: V100-A122, G129-A148, G153-M175, V188-S210, F215-L237<br>Non-cytosolic domains: M1-Q99, G149-T152, K211-R214 | TMHMMER |
| | | | | | Sugar transport proteins signature 1: L118-T134 | MOTIFS |
| 32 | 7505631CD1 | 271 | S83 S132 S232 T259 | N29 N241 | signal_cleavage: M1-A52 | SPSCAN |
| | | | | | Signal Peptide: M1-G22 | HMMER |
| | | | | | ZIP Zinc transporter: R140-M271 | HMMER_PFAM |
| | | | | | Cytosolic domains: M1-F4, H60-Q103, H199-H210, V260-M271<br>Transmembrane domains: I5-A27, L37-V59, L104-G126, L176-M198, L211-S230, V240-T259<br>Non-cytosolic domains: V28-K36, N127-Q175, K231-E239 | TMHMMER |
| 33 | 7506561CD1 | 107 | S24 S46 S98 T68 T76 | | signal_cleavage: M1-A20 | SPSCAN |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Signal Peptide: M1-A20, M1-Q21, M1-P22, M1-A25, M1-A27, M5-P22, M6-A20 | HMMER |
| | | | | | Folate receptor family: M5-L107 | HMMER_PFAM |
| | | | | | PROTEIN FOLATE RECEPTOR GLYCOPROTEIN PRECURSOR SIGNAL FOLATEBINDING MEMBRANE GPIANCHOR MULTIGENE PD006906: P22-H97 | BLAST_PRODOM |
| | | | | | FOLATE-BINDING PROTEIN DM02165\|P41439\|2-242: A2-H97 | BLAST_DOMO |
| | | | | | FOLATE-BINDING PROTEIN DM02165\|P14207\|2-254: W3-H97 | BLAST_DOMO |
| | | | | | FOLATE-BINDING PROTEIN DM02165\|P02702\|1-221: P22-H97 | BLAST_DOMO |
| | | | | | FOLATE-BINDING PROTEIN DM02165\|P15328\|22-256: A20-H97 | BLAST_DOMO |
| 34 | 7510733CD1 | 249 | S11 S176 T26 | N96 | signal_cleavage: M1-A45 | SPSCAN |
| | | | | | Major intrinsic protein: R15-Y216 | HMMER_PFAM |
| | | | | | MIP: MIP family channel proteins: S28-Y216 | HMMER_TIGRFAM |
| | | | | | Cytosolic domains: S50-R53, D135-E146, I220-M249<br>Transmembrane domains: F30-L49, F54-G76, A115-F134, P147-M169, F197-V219<br>Non-cytosolic domains: M1-E29, G77-N114, N170-N196 | TMHMMER |
| | | | | | MIP family proteins BL00221: S28-V38, Q119-D135, S165-L179, W198-G208 | BLIMPS_BLOCKS |
| | | | | | Major intrinsic protein family signature PR00783: K24-C43, G151-R173, W199-V219 | BLIMPS_PRINTS |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | TRANSMEMBRANE TRANSPORT PROTEIN AQUAPORIN INTRINSIC CHANNEL MEMBRANE WATER TONOPLAST FAMILY PD000295: E95-Y216 | BLAST_PRODOM |
| | | | | | AQUAPORIN 9 TRANSPORT TRANSMEMBRANE PD162891: V217-M249 | BLAST_PRODOM |
| | | | | | MIP FAMILY DM00228\|P47862\|15-263: L16-S79, F64-V219 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|I59266\|15-263: L16-S79, F64-V219 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|P11244\|1-253: S20-S79, A99-Y216 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|P44826\|1-251: L18-S79, A99-Y216 | BLAST_DOMO |
| 35 | 7510734CD1 | 216 | S11 S200 S201 T26 | N142 | signal_cleavage: M1-A45 | SPSCAN |
| | | | | | Major intrinsic protein: R15-H197 | HMMER_PFAM |
| | | | | | MIP: MIP family channel proteins: S28-H197 | HMMER_TIGRFAM |
| | | | | | Cytosolic domains: M1-T26, H82-K101<br>Transmembrane domains: L27-L49, T59-G81, L102-Y124<br>Non-cytosolic domains: S50-I58, Y125-V216 | TMHMMER |
| | | | | | MIP family proteins BL00221: S28-V38, V78-S88 | BLIMPS_BLOCKS |
| | | | | | Major intrinsic protein family signature PR00783: K24-C43, F64-S88, K101-V120 | BLIMPS_PRINTS |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | TRANSMEMBRANE TRANSPORT PROTEIN AQUAPORIN INTRINSIC CHANNEL MEMBRANE WATER TONOPLAST FAMILY PD000295: L18-S169 | BLAST_PRODOM |
| | | | | | MIP FAMILY DM00228\|P47862\|15-263: L16-S169 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|I59266\|15-263: L16-S169 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|P43549\|340-587: K24-S169 | BLAST_DOMO |
| | | | | | MIP FAMILY DM00228\|P11244\|1-253: S20-A163 | BLAST_DOMO |
| | | | | | MIP family signature: H82-A90 | MOTIFS |
| 36 | 7503977CD1 | 223 | S191 T93 T112 T140 T182 | N54 N116 | PROTEIN MELASTATIN CHROMOSOME TRANSMEMBRANE C05C12.3 T01H8.5 I F54D1.5 IV PD018035: K8-K209 | BLAST_PRODOM |
| 37 | 7505084CD1 | 394 | S99 S242 S394 T47 T50 T54 T127 | N243 N247 N301 | signal_cleavage: M1-A41 | SPSCAN |
| | | | | | Sodium: solute symporter family: Y58-G388 | HMMER_PFAM |
| | | | | | sss: SSS sodium solute transporter superfamily: Y58-S385 | HMMER_TIGRFAM |
| | | | | | Cytosolic domains: T47-M65, T126-G137, T203-T208, C287-K306<br>Transmembrane domains: I29-S46, V66-A88, L103-V125, G138-V160, A180-Y202, L209-G231, L269-W286, G307-V329<br>Non-cytosolic domains: M1-D28, G89-E102, D161-L179, L232-D268, S330-S394 | TMHMMER |
| | | | | | Sodium: solute symporter family proteins BL00456: Y35-G89, M111-R140, L173-G227 | BLIMPS_BLOCKS |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Sodium: solute symporter family signatures:Q170-T217 | PROFILESCAN |
| | | | | | TRANSMEMBRANE TRANSPORT PERMEASE PROTEIN SODIUM SYMPORT PROLINE COTRANSPORTER SYMPORTER GLYCOPROTEIN PD000991: Y58-I383 | BLAST_PRODOM |
| | | | | | SODIUM: SOLUTE SYMPORTER FAMILY DM00745\|P13866\|24-561:D28-P370 | BLAST_DOMO |
| | | | | | SODIUM: SOLUTE SYMPORTER FAMILY DM00745\|A53582\|24-561:D28-P370 | BLAST_DOMO |
| | | | | | SODIUM: SOLUTE SYMPORTER FAMILY DM00745\|P53790\|24-561:D28-P370 | BLAST_DOMO |
| | | | | | SODIUM: SOLUTE SYMPORTER FAMILY DM00745\|S48858\|24-561:D28-P370 | BLAST_DOMO |
| 38 | 7506950CD1 | 202 | S64 S81 T83 T157 | N38 N138 | signal_cleavage: M1-L25 | SPSCAN |
| | | | | | Signal Peptide: M10-A28 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: F42-V198 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: V68-K105, I121-N130 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: T88-F104, K120-G131 | BLIMPS_PRINTS |
| | | | | | Gamma-aminobutyric-acid receptor alpha subunit signature PR01079: T40-D51, G60-F77, F104-L116 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: R44-P164 | BLAST_PRODOM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR ALPHA2 SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL PD151117: M1-T43 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P08219\|14-456: E32-P164 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P20237\|20-556: F42-P164 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P16305\|4-443: L14-P164 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P23574\|26-465: T43-S177 | BLAST_DOMO |
| 39 | 7506951CD1 | 337 | S64 S81 S233 S310 S315 T83 T157 T190 T229 T283 T328 | N38 N138 N201 | signal_cleavage: M1-L25 | SPSCAN |
| | | | | | Signal Peptide: M10-A28 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: F42-I250 | HMMER_PFAM |
| | | | | | LIC: Cation transporter family protein: F12-S337 | HMMER_TIGRFAM |
| | | | | | Cytosolic domain: M1-K249<br>Transmembrane domain: I250-F272<br>Non-cytosolic domain: W273-S337 | TMHMMER |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: V68-K105, I121-N130, D151-Y189, Y237-S278 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion-channels signature: L146-T199 | PROFILESCAN |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: T88-F104, K120-G131, C166-C180, F244-Q256 | BLIMPS_PRINTS |
| | | | | | Gamma-aminobutyric-acid receptor alpha subunit signature PR01079: T40-D51, G60-F77, F104-L116, Y196-V208, S213-G235 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: R44-R282 | BLAST_PRODOM |
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR ALPHA2 SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL PD151117: M1-T43 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P08219\|14-456: E32-F285 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P20237\|20-556: F42-F285 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P16305\|4-443: L14-F285 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P23574\|26-465: T43-T283 | BLAST_DOMO |
| | | | | | Neurotransmitter-gated ion-channels signature: C166-C180 | MOTIFS |
| 40 | 7506954CD1 | 114 | S87 S92 T105 | N38 | signal_cleavage: M1-L25 | SPSCAN |
| | | | | | Signal Peptide: M10-A28 | HMMER |
| | | | | | Gamma-aminobutyric-acid receptor alpha subunit signature PR01079: T40-D51 | BLIMPS_PRINTS |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR ALPHA2 SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL PD151117: M1-T43 | BLAST_PRODOM |
| 41 | 7506956CD1 | 400 | S64 S81 S248 S364 T83 T157 T190 T229 T260 T287 T341 | N38 N138 N201 N326 | signal_cleavage: M1-L25 | SPSCAN |
| | | | | | Signal Peptide: M10-A28 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: F42-N251 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channel transmembrane domain: T229-W386 | HMMER_PFAM |
| | | | | | LIC: Cation transporter family protein: F12-Y389 | HMMER_TIGRFAM |
| | | | | | Cytosolic domain: T287-R371<br>Transmembrane domains: W264-F286, I372-Y389<br>Non-cytosolic domains: M1-D263, L390-P400 | TMHMMER |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: V68-K105, I121-N130, D151-Y189 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion-channels signature: L146-T199 | PROFILESCAN |
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: T88-F104, K120-G131, C166-C180 | BLIMPS_PRINTS |
| | | | | | Gamma-aminobutyric acid (GABA) receptor signature PR00253: E228-A249, M262-V283, M369-Y389 | BLIMPS_PRINTS |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Gamma-aminobutyric-acid receptor alpha subunit signature PR01079: T40-D51, G60-F77, F104-L116, Y196-V208, S213-G235, K255-V268, T359-R371, V384-V395 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: R44-T241, G235-Y285 | BLAST_PRODOM |
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL GLYCOPROTEIN PD000235: K288-P349 | BLAST_PRODOM |
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR ALPHA2 SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL PD151117: M1-T43 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P08219\|14-456: E32-T241, I225-L396 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P20237\|20-556: F42-M242, I225-T287, P353-E393 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P16305\|4-443: L14-M242, I225-E393 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P23574\|26-465: T43-M242, I225-V309, V363-L390 | BLAST_DOMO |
| | | | | | ATP/GTP-binding site motif A (P-loop): G290-S297 | MOTIFS |
| | | | | | Neurotransmitter-gated ion-channels signature: C166-C180 | MOTIFS |
| 42 | 7506959CD1 | 403 | S64 S81 S251 S367 T83 T157 T263 T290 T344 | N38 N138 N329 | signal_cleavage: M1-L25 | SPSCAN |
| | | | | | Signal Peptide: M10-A28 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel ligand binding domain: F42-N254 | HMMER_PFAM |
| | | | | | Neurotransmitter-gated ion-channel transmembrane domain: T209-W389 | HMMER_PFAM |
| | | | | | LIC: Cation transporter family protein: F12-Y392 | HMMER_TIGRFAM |
| | | | | | Cytosolic domains: N227-P232, T290-R374<br>Transmembrane domains: Y204-L226, A233-A252, W267-F289, I375-Y392<br>Non-cytosolic domains: M1-G203, R253-D266, L393-P403 | TMHMMER |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: V68-K105, I121-N130, D151-Y189, Y189-S230 | BLIMPS_BLOCKS |
| | | | | | Neurotransmitter-gated ion-channels signature: L146-R200 | PROFILESCAN |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Neurotransmitter-gated ion channel family signature PR00252: T88-F104, K120-G131, C166-C180, F196-Q208 | BLIMPS_PRINTS |
| | | | | | Gamma-aminobutyric acid (GABA) receptor signature PR00253: F205-W225, V231-A252, M265-V286, M372-Y392 | BLIMPS_PRINTS |
| | | | | | Gamma-aminobutyric-acid receptor alpha subunit signature PR01079: T40-D51, G60-F77, F104-L116, K258-V271, T362-R374, V387-V398 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: R44-K201, L182-Y288 | BLAST_PRODOM |
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL GLYCOPROTEIN PD000235: K291-P352 | BLAST_PRODOM |
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR ALPHA2 SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL PD151117: M1-T43 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P08219\|14-456: E32-T190, E188-L399 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P20237\|20-556: F42-Y189, E188-T290, P356-E396 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P16305\|4-443: L14-Y189, L155-E396 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P23574\|26-465: T43-Y189, S186-V312, V366-L393 | BLAST_DOMO |
| | | | | | ATP/GTP-binding site motif A (P-loop): G293-S300 | MOTIFS |
| | | | | | Neurotransmitter-gated ion-channels signature: C166-C180 | MOTIFS |
| 43 | 7506960CD1 | 66 | | N38 | signal_cleavage: M1-L25 | SPSCAN |
| | | | | | Signal Peptide: M10-A28 | HMMER |
| | | | | | Gamma-aminobutyric-acid receptor alpha subunit signature PR01079: T40-D51 | BLIMPS_PRINTS |
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR ALPHA2 SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL PD151117: M1-T43 | BLAST_PRODOM |
| 44 | 7510540CD1 | 89 | S42 S51 | N37 | signal_cleavage: M1-G20 | SPSCAN |
| | | | | | Cytosolic domain: A33-R89<br>Transmembrane domain: L10-N32<br>Non-cytosolic domain: M1-T9 | TMHMMER |
| | | | | | Sugar transporter signature PR00171: S21-I31 | BLIMPS_PRINTS |
| | | | | | GLUCOSE TRANSPORTER TYPE LIVER DUPLICATION TRANSMEMBRANE SUGAR TRANSPORT GLYCOPROTEIN MULTIGENE PD002509: M1-Q36 | BLAST_PRODOM |
| 45 | 7510545CD1 | 146 | S100 T8 T73 | N19 | signal_cleavage: M1-P62 | SPSCAN |
| | | | | | Signal Peptide: M46-A63 | HMMER |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Cytosolic domain: T73-G146<br>Transmembrane domain: V50-G72<br>Non-cytosolic domain: M1-S49 | TMHMMER |
| | | | | | E1-E2 ATPases phosphorylation site proteins BL00154: I57-G93, T95-V113 | BLIMPS_BLOCKS |
| | | | | | E1-E2 ATPases phosphorylation site: A81-R129 | PROFILESCAN |
| | | | | | P-type cation-transporting atpase superfamily signature PR00119: C99-V113 | BLIMPS_PRINTS |
| | | | | | Sodium/potassium-transporting ATPase signature PR00121: L92-V113 | BLIMPS_PRINTS |
| | | | | | ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATPBINDING TRANSPORT PUMP CALCIUM MAGNESIUM MEMBRANE PD000132: P36-I118 | BLAST_PRODOM |
| | | | | | E1-E2 ATPASES PHOSPHORYLATION SITE DM00115\|P18596\|43-795: E22-S49, S49-E123 | BLAST_DOMO |
| | | | | | E1-E2 ATPASES PHOSPHORYLATION SITE DM00115\|P04191\|43-795: S49-E123 | BLAST_DOMO |
| | | | | | E1-E2 ATPASES PHOSPHORYLATION SITE DM00115\|P22700\|43-795: E2-E123 | BLAST_DOMO |
| | | | | | E1-E2 ATPASES PHOSPHORYLATION SITE DM00115\|P35316\|43-799: E2-A127 | BLAST_DOMO |
| | | | | | E1-E2 ATPases phosphorylation site: D101-T107 | MOTIFS |
| 46 | 7510654CD1 | 353 | S99 T2 T205 T281 | | signal_cleavage: M1-G41 | SPSCAN |
| | | | | | Signal Peptide: M1-G41 | HMMER |
| | | | | | Sugar (and other) transporter: A29-P353 | HMMER_PFAM |
| | | | | | SP: Sugar transporter: M1-V350 | HMMER_TIGRFAM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Cytosolic domains: M1-V26, D92-K97, Y150-G155, P206-P256, D314-R319<br>Transmembrane domains: F27-P49, A69-V91, L98-A117, R127-A149, L156-V178, W183-T205, F257-A279, A294-M313, L320-H342<br>Non-cytosolic domains: A50-A68, Q118-G126, L179-R182, E280-L293, L343-P353 | TMHMMER |
| | | | | | Sugar transport proteins BL00216: L123-A172 | BLIMPS_BLOCKS |
| | | | | | Sugar transport proteins signatures: V108-L174, L293-Q344 | PROFILESCAN |
| | | | | | Sugar transporter signature PR00171: G41-I51, L124-V143, Q267-F277 | BLIMPS_PRINTS |
| | | | | | Glucose transporter signature PR00172: F257-Y278, S292-M313 | BLIMPS_PRINTS |
| | | | | | SUGAR TRANSPORT PROTEINS DM00135\|P47843\|104-456: G110-L341 | BLAST_DOMO |
| | | | | | SUGAR TRANSPORT PROTEINS DM00135\|P32037\|104-456: G110-L341 | BLAST_DOMO |
| | | | | | SUGAR TRANSPORT PROTEINS DM00135\|Q07647\|104-456: G110-L341 | BLAST_DOMO |
| | | | | | SUGAR TRANSPORT PROTEINS DM00135\|P47842\|104-456: G110-L341 | BLAST_DOMO |
| | | | | | Sugar transport proteins signature 1: G87-S104, A309-S325 | MOTIFS |
| | | | | | Sugar transport proteins signature 2: L129-R154 | MOTIFS |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| 47 | 7510660CD1 | 1155 | S205 S224 S306 S328 S612 S634 S734 S757 S809 S887 S1085 S1110 T66 T438 T567 T596 T603 T868 T871 T919 T1148 | N150 N287 N420 N502 N1058 | haloacid dehalogenase-like hydrolase: V527-A843 | HMMER_PFAM |
| | | | | | Cytosolic domains: M1-T66, V122-T444, L1030-P1054, S1110-S1155<br>Transmembrane domains: V67-W89, A99-S121, F445-I464, I1007-Y1029, S1055-Y1077, L1092-G1109<br>Non-cytosolic domains: G90-E98, E465-N1006, K1078-P1091 | TMHMMER |
| | | | | | E1-E2 ATPases phosphorylation site proteins BL00154: V489-G525, V527-V545, E681-F721, T817-L840, A912-M945 | BLIMPS_BLOCKS |
| | | | | | P-type cation-transporting atpase superfamily signature PR00119: D348-E362, C531-V545, T697-D707, C820-L839 | BLIMPS_PRINTS |
| | | | | | PROBABLE CALCIUMTRANSPORTING ATPASE HYDROLASE CALCIUM TRANSPORT TRANSMEMBRANE PHOSPHORYLATION MAGNESIUM ATPBINDING PD023991: D901-G1147 | BLAST_PRODOM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | ATPASE PROBABLE CALCIUMTRANSPORTING HYDROLASE CALCIUM TRANSPORT TRANSMEMBRANE PHOSPHORYLATION MAGNESIUM ATPBINDING PD150086: G208-V267 | BLAST_PRODOM |
| | | | | | ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATPBINDING TRANSPORT PUMP CALCIUM MAGNESIUM MEMBRANE PD000132: I312-V548, I674-L719 | BLAST_PRODOM |
| | | | | | E1-E2 ATPASES PHOSPHORYLATION SITE DM00115\|P37367\|60-746: I312-V390, V414-G550, D581-L839 | BLAST_DOMO |
| | | | | | E1-E2 ATPASES PHOSPHORYLATION SITE DM00115\|A46284\|47-821: E316-E566, L595-L839 | BLAST_DOMO |
| | | | | | E1-E2 ATPASES PHOSPHORYLATION SITE DM00115\|S27763\|47-821: E316-E566, L595-L839 | BLAST_DOMO |
| | | | | | E1-E2 ATPASES PHOSPHORYLATION SITE DM00115\|P39168\|83-733: W309-V391, T388-L839 | BLAST_DOMO |
| | | | | | E1-E2 ATPases phosphorylation site: D533-T539 | MOTIFS |
| 48 | 7510661CD1 | 606 | S205 S224 S306 S328 T66 T438 T567 T596 | N150 N287 N420 N502 | signal_cleavage: M1-G46 | SPSCAN |
| | | | | | Cytosolic domains: M1-T66, V122-T444<br>Transmembrane domains: V67-W89, A99-S121, F445-I464<br>Non-cytosolic domains: G90-E98, E465-P606 | TMHMMER |
| | | | | | E1-E2 ATPases phosphorylation site proteins BL00154: V489-G525, V527-V545 | BLIMPS_BLOCKS |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | E1-E2 ATPases phosphorylation site: I521-S561 | PROFILESCAN |
| | | | | | ATPASE PROBABLE CALCIUMTRANSPORTING HYDROLASE CALCIUM TRANSPORT TRANSMEMBRANE PHOSPHORYLATION MAGNESIUM ATPBINDING PD150086: G208-V267 | BLAST_PRODOM |
| | | | | | ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATPBINDING TRANSPORT PUMP CALCIUM MAGNESIUM MEMBRANE PD000132: I312-V548 | BLAST_PRODOM |
| | | | | | E1-E2 ATPases phosphorylation site: D533-T539 | MOTIFS |
| 49 | 7510680CD1 | 462 | S13 S18 S225 S314 S373 T33 T323 T351 | N229 N249 | 2A0119: cation transport protein: M1-Q460 | HMMER_TIGRFAM |
| | | | | | Cytosolic domains: M1-I48, D109-S120, R202-Q283, R370-T381, I451-K462<br>Transmembrane domains: A49-V71, V86-A108, F121-L143, V179-W201, I284-L306, M347-G369, A382-L404, I428-P450<br>Non-cytosolic domains: S72-Q85, R144-Q178, E307-T346, R405-S427 | TMHMMER |
| | | | | | SUGAR TRANSPORT PROTEINS DM00032\|P30638\|80-152: R45-K115 | BLAST_DOMO |
| | | | | | do VESICLE; SYNAPTIC; SV2; FORM; DM08835\|S34961\|180-344: I119-N249 | BLAST_DOMO |
| 50 | 7505145CD1 | 366 | S271 S287 S317 T10 T136 Y61 | N156 | 2_A_01_02: Multidrug resistance protein: G90-T225 | HMMER_TIGRFAM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Cytosolic domains: P47-S85, T136-Q175, P223-R274 Transmembrane domains: V24-L46, V86-G108, P118-A135, G176-L198, A203-L222, L275-T297 Non-cytosolic domains: M1-R23, A109-R117, P199-M202, H298-A366 | TMHMMER |
| | | | | | Tetracycline resistance protein signature PR01035: L34-L50, M177-P199, A203-P223 | BLIMPS_PRINTS |
| | | | | | TETRACYCLINE TRANSPORTERLIKE PROTEIN TRANSPORT MRNA PD023345: L226-S304 | BLAST_PRODOM |
| | | | | | TETRACYCLINE TRANSPORTERLIKE PROTEIN TRANSPORT MRNA PD029169: E51-G89 | BLAST_PRODOM |
| | | | | | TETRACYCLINE TRANSPORTERLIKE PROTEIN TRANSPORT MRNA PD025998: M1-F29 | BLAST_PRODOM |
| 51 | 7505162CD1 | 295 | S37 S75 S164 S210 T216 | | 2A0104: phosphoglycerate transporter protein: F18-V285 | HMMER_TIGRFAM |
| | | | | | Cytosolic domains: M1-Y6, D72-R77, T157-T168, K240-Q295 Transmembrane domains: G7-F26, L49-S71, W78-V100, F134-A156, L169-H188, L220-V239 Non-cytosolic domains: N27-D48, P101-Q133, N189-E219 | TMHMMER |
| | | | | | glpT family of transporters proteins BL00942: M17-K29, L44-L86, W128-L147, L169-E205, S210-F250 | BLIMPS_BLOCKS |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | GLPT FAMILY OF TRANSPORTERS DM02439\|P09836\|1-401: Y6-W281 | BLAST_DOMO |
| | | | | | GLPT FAMILY OF TRANSPORTERS DM02439\|P37948\|1-403: Y6-S258 | BLAST_DOMO |
| | | | | | GLPT FAMILY OF TRANSPORTERS DM02439\|P08194\|1-403: Y6-K255 | BLAST_DOMO |
| | | | | | GLPT FAMILY OF TRANSPORTERS DM02439\|P12681\|1-405: Y6-T216 | BLAST_DOMO |
| 52 | 7505469CD1 | 229 | S22 S66 Y197 | N18 | Cytosolic domains: M1-E44, E100-S111, K186-Y197<br>Transmembrane domains: I45-P67, Y77-A99, Y112-V134, Y163-V185, A198-P220<br>Non-cytosolic domains: K68-S76, V135-P162, W221-L229 | TMHMMER |
| | | | | | Amino acid permeases proteins BL00218: L48-S76, S80-S111, V180-V208 | BLIMPS_BLOCKS |
| | | | | | MYELOBLAST KIAA0245 PD078048: M1-Q40 | BLAST_PRODOM |
| | | | | | dò ANTIPORTER; ORNITHINE; PUTRESCINE; TRANSPORT; DM01125\|P45539\|1-318: T38-L211 | BLAST_DOMO |
| 53 | 7505475CD1 | 637 | S30 S50 S134 S249 S353 S491 T111 T206 T558 T572 T624 | N71 N84 N91 N109 N130 N241 N436 N544 N576 | Signal Peptide: M26-A45, M26-M51 | HMMER |
| | | | | | ABC transporter: G507-K637 | HMMER_PFAM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Cytosolic domains: M1-E29, K244-F263, S319-K324, D418-K637<br>Transmembrane domains: S30-S49, E221-T243, W264-I286, G296-L318, A325-F347, T395-F417<br>Non-cytosolic domains: S50-N220, T287-T295, Y348-Y394 | TMHMMER |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P41233\|839-1045: I478-I635 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P34358\|611-816: I478-I635 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P41233\|1851-2058: I478-I635 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|P44785\|2-216: I478-I635 | BLAST_DOMO |
| | | | | | ATP/GTP-binding site motif A (P-loop): G514-S521 | MOTIFS |
| 54 | 7505568CD1 | 90 | | N74 | signal_cleavage: M1-G66 | SPSCAN |
| | | | | | Cytosolic domain: M1-S55<br>Transmembrane domain: L56-V78<br>Non-cytosolic domain: N79-Q90 | TMHMMER |
| | | | | | Sugar transporter signature PR00171: S68-V78 | BLIMPS_PRINTS |
| 55 | 7506953CD1 | 327 | S109 S175 S291 T66 T105 T187 T214 T268 | N38 N77 N253 | signal_cleavage: M1-L25 | SPSCAN |
| | | | | | Signal Peptide: M10-A28 | HMMER |
| | | | | | Neurotransmitter-gated ion-channel transmembrane domain: T133-W313 | HMMER_PFAM |
| | | | | | LIC: Cation transporter family protein: F12-Y316 | HMMER_TIGRFAM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Cytosolic domains: M1-K125, T214-R298<br>Transmembrane domains: I126-F148, W191-F213, I299-Y316<br>Non-cytosolic domains: W149-D190, L317-P327 | TMHMMER |
| | | | | | Neurotransmitter-gated ion-channels proteins BL00236: Y113-S154 | BLIMPS_BLOCKS |
| | | | | | Gamma-aminobutyric acid (GABA) receptor signature PR00253: F129-W149, V155-A176, M189-V210, M296-Y316 | BLIMPS_PRINTS |
| | | | | | Gamma-aminobutyric-acid receptor alpha subunit signature PR01079: Y72-V84, S89-G111, K182-V195, T286-R298, V311-V322 | BLIMPS_PRINTS |
| | | | | | CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: Y65-Y212 | BLAST_PRODOM |
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL GLYCOPROTEIN PD000235: K215-P276 | BLAST_PRODOM |
| | | | | | GAMMAAMINOBUTYRICACID RECEPTOR ALPHA2 SUBUNIT PRECURSOR GABAA POSTSYNAPTIC MEMBRANE IONIC CHANNEL PD151117: M1-T43 | BLAST_PRODOM |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P08219\|14-456: E32-T66, A64-L323 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P20237\|20-556: R44-T214, P280-E320 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P16305\|4-443: R44-E320 | BLAST_DOMO |
| | | | | | NEUROTRANSMITTER-GATED ION-CHANNELS DM00560\|P23574\|26-465: T43-V84, Y65-V236, V290-L317 | BLAST_DOMO |
| | | | | | ATP/GTP-binding site motif A (P-loop): G217-S224 | MOTIFS |
| 56 | 7510176CD1 | 40 | | | signal_cleavage: M1-T32 | SPSCAN |
| | | | | | Cytosolic domain: M1-R12<br>Transmembrane domain: F13-I35<br>Non-cytosolic domain: L36-R40 | TMHMMER |
| 57 | 7510541CD1 | 104 | S41 T60 | N74 | Cytosolic domain: H101-L104<br>Transmembrane domain: A78-L100<br>Non-cytosolic domain: M1-N77 | TMHMMER |
| | | | | | TRANSPORTER PROTEIN PD182518: M1-M70 | BLAST_PRODOM |
| 58 | 7510923CD1 | 296 | S41 S254 T60 Y134 | N74 N247 N248 N252 | Cytosolic domains: A90-G95, K165-D184, K234-Q296<br>Transmembrane domains: S67-Y89, I96-L118, L143-I164, W185-M207, G211-Y233<br>Non-cytosolic domains: M1-T66, K119-K142, R208-L210 | TMHMMER |
| | | | | | TRANSPORTER PROTEIN PD182518: M1-M70 | BLAST_PRODOM |
| | | | | | TRANSPORTER PROTEIN PD138403: Y233-Q285 | BLAST_PRODOM |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | ACID AMINO PROTEIN TRANSPORTER PERMEASE TRANSMEMBRANE INTERGENIC REGION PUTATIVE PROLINE PD001875: D61-Y233 | BLAST_PRODOM |
| 59 | 7510984CD1 | 1364 | S55 S347 S408 S487 S532 S615 S742 S774 S858 S961 S980 S1138 S1342 T151 T200 T304 T520 T524 T525 T600 T700 T758 T760 T783 T888 T913 T943 T945 T950 T1213 T1321 | N10 N406 N698 N781 N829 N985 N1050 | Cytosolic domains: S53-N72, H125-L135, R191-L301, Q369-Q427, A478-A537, R598-V1063, Q1179-E1249, N1296-Y1364<br>Transmembrane domains: A30-G52, L73-S95, H105-Y124, L136-L158, R168-I190, S302-V324, F351-L368, L428-I450, I455-V477, I538-V560, V575-V597, Y1064-V1086, V1156-I1178, V1250-L1272, L1276-L1295<br>Non-cytosolic domains: M1-D29, D96-L104, D159-L167, D325-E350, L451-Y454, G561-S574, E1087-A1155, H1273-E1275 | TMHMMER |
| | | | | | ABC transporter transmembrane region.: L1012-V1299, L318-L590 | HMMER_PFAM |
| | | | | | ABC transporter: G706-G906 | HMMER_PFAM |
| | | | | | ATPases associated with a variety of cellular proteins: R705-L914 | HMMER_SMART |
| | | | | | ABC transporters family proteins BL00211: I711-L722, L830-D861 | BLIMPS_BLOCKS |
| | | | | | ABC transporters family signature: I812-D861 | PROFILESCAN |
| | | | | | Presenilin 1 signature PR01073: D964-E975 | BLIMPS_PRINTS |
| | | | | | Sulphonylurea receptor family signature PR01092: G25-G52, W65-F79, V122-L136, V204-K227, V357-V379 | BLIMPS_PRINTS |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | Sulphonylurea receptor type 1 family signature PR01093: E9-Q19, A269-D277, A617-P629, Y638-E654, A1047-C1057 | BLIMPS_PRINTS |
| | | | | | SULFONYLUREA RECEPTOR ATPBINDING TRANSPORT 2B 2A TRANSMEMBRANE PHOSPHORYLATION GLYCOPROTEIN 1B PD005449: M1-R297 | BLAST_PRODOM |
| | | | | | SULFONYLUREA RECEPTOR ATPBINDING TRANSPORT TRANSMEMBRANE PHOSPHORYLATION GLYCOPROTEIN 1B ALTERNATIVE SPLICING PD011659: F591-T695 | BLAST_PRODOM |
| | | | | | SULFONYLUREA RECEPTOR ATPBINDING TRANSPORT 2B 2A TRANSMEMBRANE PHOSPHORYLATION GLYCOPROTEIN 1B PD005248: T913-I1010 | BLAST_PRODOM |
| | | | | | SULFONYLUREA RECEPTOR ATPBINDING TRANSPORT TRANSMEMBRANE PHOSPHORYLATION GLYCOPROTEIN 1B PD151487: R298-Y356 | BLAST_PRODOM |
| | | | | | do RESISTANCE; MULTIDRUG; SAUROLEISHMANIA; C3F10.11C; DM01525\|Q09427\|906-1342: T907-L1334 | BLAST_DOMO |
| | | | | | do RESISTANCE; MULTIDRUG; SAUROLEISHMANIA; C3F10.11C; DM01742\|Q09427\|213-630: G214-Q632 | BLAST_DOMO |
| | | | | | ABC TRANSPORTERS FAMILY DM00008\|Q09427\|690-904: D691-G906 | BLAST_DOMO |

# Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|---|---|---|---|---|---|---|
| | | | | | do RESISTANCE; MULTIDRUG; SAUROLEISHMANIA; C3F10.11C; DM01525\|P33527\|868-1291: L967-S1325 | BLAST_DOMO |
| | | | | | ABC transporters family signature: L830-L844 | MOTIFS |
| | | | | | ATP/GTP-binding site motif A (P-loop): G713-S720 | MOTIFS |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 60/7509332CB1/895 | 1-895, 4-444, 26-450, 55-264, 133-264, 133-505, 133-849, 229-505, 354-495, 406-503 |
| 61/7509102CB1/1623 | 1-932, 1-986, 1-1623, 447-748, 451-653, 630-1189, 981-1215, 981-1326, 981-1342, 981-1379, 981-1419, 981-1425, 984-1395, 1515-1611 |
| 62/7509132CB1/1802 | 1-477, 39-850, 61-1802, 581-1046, 659-889, 662-1278, 662-1286, 712-1634, 716-1634, 736-1280, 750-1240, 750-1634, 802-1634, 816-1055, 818-923, 825-1468, 862-1280, 930-1634, 958-1233, 959-1634, 961-1236, 963-1524, 1086-1327, 1123-1418, 1139-1634, 1178-1455, 1182-1262, 1182-1409, 1182-1423, 1182-1436, 1373-1609, 1373-1770, 1412-1649 |
| 63/7509136CB1/2139 | 1-177, 1-250, 1-374, 1-457, 1-468, 1-495, 1-2139, 77-667, 77-690, 77-693, 77-841, 78-705, 111-316, 216-333, 242-443, 357-459, 559-1332, 703-1007, 703-1176, 703-1332, 704-1330, 705-1332, 721-1332, 723-1239, 730-943, 763-1332, 781-1320, 787-1364, 796-954, 804-1373, 804-1418, 807-895, 815-1062, 817-1359, 827-1365, 838-1000, 842-1017, 866-1389, 871-1315, 888-1385, 901-1397, 936-1497, 956-1223, 973-1509, 976-1303, 977-1526, 977-1565, 985-1510, 986-1521, 987-1843, 998-1326, 1002-1542, 1022-1717, 1033-1563, 1060-1537, 1080-1689, 1081-1655, 1082-1608, 1096-1612, 1106-1719, 1121-1682, 1130-1652, 1145-1666, 1184-1446, 1189-1675, 1202-1462, 1202-1566, 1205-1577, 1207-1502, 1207-1552, 1214-1552, 1305-1802, 1320-1800, 1322-1775, 1338-1633, 1338-1635, 1339-1801, 1354-1848, 1371-2110, 1399-1552, 1419-1641, 1419-1661, 1445-2098, 1451-1709, 1459-1753, 1470-1768, 1477-2105, 1567-1672, 1567-1674, 1567-1681, 1567-1687, 1567-1688, 1572-2139, 1688-2139, 1741-2105, 1808-2010, 1808-2134, 1861-2105, 1986-2093, 1986-2105, 1991-2105, 1992-2105, 1996-2105 |
| 64/7509178CB1/1461 | 1-252, 1-703, 1-913, 1-1459, 24-252, 248-804, 288-523, 351-1163, 361-1164, 419-1164, 420-1164, 444-1164, 452-699, 463-1098, 465-1165, 505-1163, 506-1034, 506-1036, 509-1164, 523-1164, 528-830, 528-897, 540-1164, 570-1420, 571-1461, 575-1461, 578-1112, 594-1461, 601-1434, 602-1461, 620-1461, 629-1460, 653-914, 670-1173, 672-1164, 684-1190, 687-906, 688-906, 716-1160, 716-1359, 716-1376, 726-1321, 773-1289, 798-1461, 799-1461, 811-1343, 833-1106, 833-1428, 953-1395, 1037-1223, 1144-1293, 1204-1459 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 65/7509214CB1/738 | 1-562, 1-738, 5-116, 41-139, 135-289, 135-305, 135-324, 135-359, 135-382, 135-386, 135-415, 135-430, 135-579, 135-623, 135-727, 136-386, 138-561, 141-415, 142-393, 146-634, 146-738, 147-612, 149-384, 149-394, 149-434, 150-372, 151-405, 151-417, 151-440, 152-576, 164-324, 165-383, 166-359, 166-738, 167-324, 170-398, 184-296, 186-442, 187-398, 187-433, 190-463, 203-434, 203-651, 204-468, 204-738, 205-470, 205-475, 209-482, 209-738, 213-481, 214-350, 216-475, 217-468, 221-738, 222-725, 230-506, 231-720, 233-728, 234-737, 238-496, 245-540, 246-479, 246-728, 249-736, 250-536, 254-723, 256-705, 256-730, 261-722, 261-723, 263-724, 269-738, 270-719, 271-505, 272-525, 273-726, 275-728, 277-725, 278-449, 278-550, 280-485, 280-738, 281-724, 282-508, 282-738, 283-542, 283-722, 283-725, 283-728, 284-699, 284-728, 286-580, 289-533, 289-554, 289-643, 291-556, 292-721, 296-722, 296-738, 298-724, 302-730, 303-721, 304-738, 305-720, 308-721, 308-722, 309-542, 309-722, 309-728, 310-734, 310-738, 311-588, 312-736, 313-738, 314-385, 314-573, 315-571, 318-723, 320-553, 322-721, 324-637, |
| | 325-722, 327-732, 328-592, 328-709, 328-738, 329-722, 330-727, 330-728, 331-738, 334-723, 335-723, 335-725, 337-723, 338-722, 341-606, 342-573, 342-640, 344-510, 345-725, 345-734, 347-581, 347-660, 347-726, 348-719, 350-721, 352-721, 352-738, 354-654, 355-651, 358-588, 358-703, 360-687, 361-726, 363-724, 364-716, 366-722, 369-725, 370-726, 371-581, 374-703, 380-718, 381-589, 381-722, 382-720, 382-723, 382-725, 384-725, 385-661, 385-728, 389-638, 390-539, 390-618, 393-725, 393-738, 409-720, 413-722, 413-723, 414-696, 414-737, 418-722, 421-726, 426-676, 426-683, 426-685, 426-711, 426-722, 427-721, 427-728, 428-696, 428-719, 429-645, 429-736, 431-726, 431-736, 432-722, 432-723, 432-724, 433-728, 434-722, 434-727, 436-722, 437-722, 437-725, 440-738, 443-722, 444-713, 446-728, 447-722, 447-725, 448-718, 448-721, 448-725, 448-728, 449-724, 450-722, 462-727, 462-728, 467-722, 472-728, 473-722, 474-737, 476-711, 477-721, 478-728, 479-722, 479-723, 482-721, 483-719, 487-722, 489-720, 497-724, 498-723, 501-722, 501-737, 502-738, 503-722, 505-727, 510-722, 511-723, 512-728, |
| | 599-731, 610-738, 611-722, 628-728, 629-738, 648-722, 671-722 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 66/7509244CB1/2106 | 1-177, 1-250, 1-374, 1-457, 1-468, 1-495, 1-2106, 77-667, 77-690, 77-693, 77-819, 77-820, 77-832, 77-838, 77-871, 77-878, 77-886, 77-893, 77-896, 77-980, 77-988, 77-999, 77-1002, 78-714, 78-758, 78-857, 78-858, 78-896, 111-316, 216-333, 241-888, 242-443, 508-1299, 705-737, 726-994, 989-1509, 996-1684, 1000-1530, 1027-1504, 1047-1656, 1048-1622, 1049-1575, 1063-1579, 1073-1686, 1088-1649, 1097-1619, 1112-1633, 1151-1413, 1156-1642, 1169-1429, 1169-1533, 1172-1544, 1174-1469, 1174-1519, 1181-1519, 1272-1769, 1287-1767, 1289-1742, 1305-1600, 1305-1602, 1306-1768, 1321-1815, 1338-2077, 1366-1519, 1386-1608, 1386-1628, 1412-2065, 1418-1676, 1426-1720, 1437-1735, 1444-2072, 1534-1639, 1534-1641, 1534-1648, 1534-1654, 1534-1655, 1539-2106, 1655-2106, 1708-2072, 1775-1977, 1775-2101, 1828-2072, 1953-2060, 1953-2072, 1958-2072, 1959-2072, 1963-2072 |
| 67/7509256CB1/2334 | 1-236, 1-2334, 18-115, 18-142, 18-143, 18-178, 18-255, 18-281, 18-339, 18-352, 18-394, 18-428, 18-480, 18-487, 18-644, 19-199, 22-357, 51-190, 51-654, 54-653, 72-204, 85-366, 85-373, 85-593, 85-690, 85-842, 86-636, 86-719, 86-826, 90-610, 97-357, 104-733, 127-448, 127-877, 127-887, 127-912, 128-672, 128-825, 128-826, 128-861, 128-876, 128-888, 128-889, 128-894, 128-906, 128-936, 128-949, 128-950, 128-983, 128-991, 128-992, 128-993, 128-997, 128-1008, 128-1030, 128-1037, 128-1039, 128-1058, 128-1066, 128-1070, 138-884, 138-907, 138-949, 138-977, 138-1058, 138-1087, 138-1102, 189-713, 190-543, 222-728, 255-980, 262-453, 354-997, 363-638, 412-1081, 421-1122, 444-1125, 450-886, 456-733, 469-872, 527-955, 529-627, 535-1525, 556-1527, 557-1000, 585-1152, 585-1154, 587-1203, 590-1148, 596-816, 599-918, 599-1173, 613-856, 626-1236, 639-1020, 646-1206, 660-1526, 662-1213, 662-1525, 670-1230, 671-1525, 675-1150, 676-1230, 679-1193, 701-1005, 715-1526, 721-1526, 727-1527, 732-1526, 733-1525, 734-1106, 734-1481, 736-1527, 738-1525, 745-1527, 750-1051, 773-1020, 775-1527, |
| | 785-864, 785-1202, 792-1525, 805-1525, 806-1356, 808-1247, 840-1067, 861-1241, 881-1571, 884-1391, 897-1459, 927-1525, 929-1462, 948-1561, 963-1106, 988-1071, 994-1328, 1023-1569, 1029-1300, 1056-1631, 1070-1297, 1075-1917, 1080-1696, 1085-1709, 1100-1262, 1110-1349, 1116-1378, 1116-1381, 1140-1356, 1145-1405, 1147-1455, 1150-1870, 1164-1743, 1176-1878, 1185-1856, 1211-1801, 1215-1448, 1220-1534, 1235-1506, 1235-1626, 1246-1766, 1253-1800, 1260-1875, 1266-1549, 1285-1532, 1351-1451, 1352-1882, 1353-2170, 1356-2170, 1366-1642, 1386-1669, 1400-1808, 1407-2170, 1408-2170, 1409-1647, 1415-1888, 1416-2170, 1418-2169, 1418-2170, 1423-2167, 1454-1885, 1456-2172, 1468-1575, 1475-2170, 1486-1791, 1504-1928, 1508-1928, 1523-1928, 1524-2020, 1526-1928, 1527-1925, 1547-1926, 1555-1776, 1558-1888, 1562-1849, 1562-1927, 1595-1925, 1600-1859, 1682-1919, 1708-1925, 1714-1927, 1766-1928, 1807-2076, 1839-2170, 1849-1925, 1862-2330, 1928-2334 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 68/7509395CB1/1475 | 1-717, 28-927, 30-1473, 53-266, 267-818, 302-537, 365-1177, 375-1178, 433-1178, 434-1178, 458-1178, 466-713, 477-1112, 479-1179, 519-1177, 520-1048, 520-1050, 523-1178, 537-1178, 542-844, 542-911, 554-1178, 584-1434, 585-1475, 589-1475, 592-1126, 608-1475, 615-1448, 616-1475, 634-1475, 643-1474, 667-928, 684-1187, 686-1178, 698-1204, 701-920, 702-920, 730-1174, 730-1373, 730-1390, 740-1335, 787-1303, 812-1475, 813-1475, 825-1357, 847-1120, 847-1442, 967-1409, 1051-1237, 1158-1307, 1218-1473 |
| 69/7503287CB1/1295 | 1-278, 1-423, 1-552, 1-1295, 17-667, 19-623, 35-667, 37-471, 37-593, 37-709, 37-763, 37-816, 37-870, 37-877, 37-880, 37-883, 37-885, 37-886, 63-669, 76-886, 81-884, 89-886, 105-886, 106-886, 126-495, 126-662, 126-768, 126-771, 128-768, 149-886, 277-644, 284-768, 800-976, 800-1044, 800-1295, 869-1146, 911-1164, 921-1166, 921-1175, 955-1230 |
| 70/7503320CB1/1386 | 1-575, 1-758, 1-855, 1-1386, 146-998, 152-997, 214-997, 236-865, 237-997, 238-998, 239-997, 243-997, 249-1070, 250-514, 252-997, 280-998, 283-997, 399-641, 407-643, 413-998, 530-775, 531-703, 572-793, 572-795, 587-1222, 668-1222, 702-927, 723-1149, 724-1261, 731-1240, 778-1230, 799-1376, 811-1230, 836-953, 836-1248, 865-970, 879-1169, 891-1312, 893-1353, 904-1359, 905-1386, 916-1365, 984-1222, 992-1222, 998-1240, 1002-1204, 1051-1219, 1118-1365, 1163-1376, 1172-1366 |
| 71/7503335CB1/2213 | 1-323, 5-794, 7-298, 7-452, 16-607, 20-216, 20-2008, 21-284, 32-607, 33-309, 33-369, 33-373, 33-457, 33-539, 33-613, 33-712, 33-729, 33-754, 35-284, 39-626, 43-520, 43-669, 68-699, 71-707, 74-306, 74-331, 74-367, 77-530, 95-489, 113-608, 142-802, 189-553, 241-761, 243-794, 260-696, 269-529, 277-839, 286-597, 299-810, 317-783, 357-547, 397-517, 412-913, 612-690, 700-942, 984-1650, 989-1538, 997-1318, 1005-1288, 1010-1262, 1014-1610, 1014-1705, 1074-1302, 1115-1342, 1123-1392, 1151-1420, 1153-1638, 1164-1436, 1167-1459, 1167-1500, 1167-1833, 1169-1971, 1188-1400, 1224-1482, 1249-1438, 1274-1957, 1280-1786, 1287-1744, 1287-1961, 1295-1597, 1295-1804, 1323-1899, 1365-1992, 1371-1650, 1382-2213, 1389-1945, 1394-1852, 1399-2025, 1408-1564, 1439-1966, 1467-1948, 1481-1728, 1481-1907, 1509-1664, 1510-2025, 1613-1733 |
| 72/7503952CB1/1289 | 1-1289, 295-852, 354-893 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 73/7504530CB1/1358 | 1-356, 3-295, 3-688, 3-690, 3-698, 3-731, 3-863, 17-242, 17-290, 17-457, 17-459, 31-356, 33-356, 33-1358, 36-356, 38-196, 45-503, 50-346, 60-295, 60-513, 60-668, 60-859, 62-363, 73-356, 75-835, 80-229, 80-364, 93-353, 101-356, 109-353, 109-356, 111-356, 206-699, 206-915, 206-945, 206-946, 206-948, 231-356, 231-363, 231-469, 234-1192, 249-948, 259-750, 302-1191, 333-948, 334-1191, 365-858, 385-1191, 390-1191, 407-518, 412-544, 421-670, 435-713, 476-1192, 534-849, 591-1191, 602-1178, 610-1139, 614-1229, 622-866, 622-948, 669-1230, 691-904, 706-958, 706-960, 733-1230, 765-1226, 766-1230, 826-1211, 867-1221, 926-1230, 927-1125, 932-1230, 947-1221, 947-1230, 993-1264, 1000-1272, 1058-1230, 1146-1358 |
| 74/7509303CB1/2232 | 1-634, 15-2217, 29-436, 29-673, 29-679, 29-682, 29-683, 29-690, 29-697, 29-713, 29-891, 33-644, 61-341, 209-718, 212-491, 223-462, 255-700, 260-671, 335-557, 616-718, 718-1010, 719-1022, 719-1271, 719-1279, 719-1294, 719-1339, 719-1359, 728-1316, 729-1202, 746-1386, 747-1246, 750-882, 750-983, 756-1143, 770-1398, 772-1263, 774-1325, 779-1417, 788-1460, 798-1314, 801-1414, 805-1404, 823-938, 826-1376, 837-1389, 840-1461, 846-1401, 857-1507, 864-1383, 886-1560, 893-1500, 895-1419, 895-1601, 898-1376, 900-1398, 902-1544, 909-1372, 910-1540, 912-1565, 917-1563, 918-1565, 920-1195, 923-1544, 928-1581, 932-1526, 932-1546, 936-1437, 938-1904, 943-1159, 953-1904, 960-1385, 965-1904, 967-1406, 967-1536, 983-1904, 985-1566, 991-1255, 991-1637, 992-1221, 992-1255, 1001-1490, 1004-1568, 1004-1616, 1006-1904, 1007-1904, 1008-1904, 1017-1262, 1017-1372, 1021-1638, 1022-1387, 1025-1315, 1026-1387, 1033-1630, 1038-1361, 1038-1581, 1043-1904, 1044-1439, 1044-1904, 1049-1302, 1051-1417, 1053-1457, 1053-1527, 1053-1602, 1053-1680, 1061-1397, 1067-1703, |
| | 1068-1694, 1085-1588, 1087-1584, 1088-1588, 1095-1904, 1097-1627, 1098-1716, 1099-1741, 1104-1419, 1107-1904, 1108-1598, 1108-1700, 1108-1724, 1108-1785, 1108-1904, 1109-1387, 1110-1723, 1111-1651, 1123-1904, 1131-1549, 1131-1683, 1149-1657, 1158-1604, 1158-1716, 1161-1734, 1162-1731, 1175-1744, 1195-1733, 1195-1750, 1198-1904, 1199-1631, 1203-1876, 1203-1904, 1214-1904, 1218-1584, 1222-1309, 1229-1484, 1233-1673, 1281-1523, 1286-1496, 1290-1578, 1307-1674, 1311-1594, 1315-1420, 1315-1582, 1315-1672, 1315-1874, 1315-1928, 1325-1988, 1327-1665, 1360-1661, 1367-1904, 1367-2042, 1370-1574, 1379-1665, 1382-1530, 1382-1624, 1384-1732, 1385-1529, 1390-1592, 1391-1667, 1393-1961, 1400-1930, 1401-1897, 1407-1657, 1407-1665, 1435-1904, 1440-1678, 1440-1734, 1440-2015, 1441-1656, 1443-1933, 1443-1992, 1444-1989, 1445-2095, 1451-1615, 1451-1663, 1451-1960, 1453-1971, 1463-1705, 1466-2061, 1467-2023, 1471-1731, 1480-1761, 1480-2123, 1483-2116, 1484-1938, 1485-2123, 1487-2013, 1501-1785, 1511-2065, 1526-1773, 1542-2081, 1543-2033, 1550-1680, |

Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1555-1939, 1558-2143, 1566-2139, 1568-2204, 1569-2189, 1573-2200, 1577-2186, 1587-2199, 1591-2197, 1593-2168, 1594-2191, 1606-2181, 1611-2205, 1613-2217, 1617-2088, 1664-2139, 1678-2217, 1681-2027, 1689-2158, 1703-2100, 1704-1948, 1704-1949, 1704-1963, 1704-2069, 1704-2113, 1724-2206, 1725-1986, 1727-2021, 1759-2139, 1768-1971, 1769-2211, 1786-2097, 1791-2021, 1791-2130, 1796-2056, 1796-2159, 1813-1995, 1816-2130, 1820-2072, 1821-1994, 1822-2071, 1825-2058, 1831-2040, 1832-2058, 1832-2108, 1851-2232, 1856-2118, 1867-2129, 1871-2133, 1871-2154, 1881-2080, 1891-2168, 1910-2168, 1918-2165, 1927-2218, 1941-2186, 1941-2232, 1946-2231, 1953-2209, 1980-2200, 1980-2218, 1981-2231 |
| 75/7509910CB1/2230 | 1-323, 5-796, 7-298, 7-413, 10-822, 16-607, 20-216, 20-2230, 21-284, 32-607, 33-309, 33-369, 33-373, 33-457, 33-539, 33-613, 33-712, 33-729, 33-754, 40-808, 43-520, 43-669, 71-707, 74-306, 74-331, 74-367, 77-530, 87-531, 95-489, 142-803, 189-553, 260-696, 267-1048, 269-529, 286-597, 299-811, 317-783, 337-797, 357-547, 397-517, 518-1226, 522-1226, 522-1230, 528-1230, 549-1300, 570-1226, 581-1300, 587-932, 587-1215, 588-1226, 605-1215, 668-1215, 689-1226, 725-1230, 742-1229, 783-1300, 793-1226, 896-1300, 962-1732, 965-1432, 982-1585, 1021-1580, 1048-1863, 1071-1392, 1078-1600, 1102-1392, 1114-1870, 1114-1872, 1124-1863, 1124-1872, 1159-1478, 1211-1760, 1219-1540, 1227-1510, 1232-1484, 1236-1832, 1236-1854, 1296-1524, 1337-1564, 1345-1614, 1373-1642, 1375-1860, 1386-1658, 1389-1681, 1389-1722, 1410-1622, 1446-1704, 1471-1660, 1496-2179, 1502-2008, 1509-1966, 1509-2183, 1517-1819, 1517-1958, 1545-2121, 1593-1872, 1601-1955, 1611-1955, 1611-2167, 1612-1955, 1616-2074, 1630-1786, 1638-1955, 1661-2188, 1689-2170, 1703-1950, 1703-1991, 1732-2230, 1835-1955 |
| 76/7509982CB1/5966 | 1-273, 1-5966, 7-483, 24-272, 28-299, 42-317, 42-524, 42-551, 42-607, 42-623, 42-638, 42-643, 42-674, 42-675, 42-680, 42-681, 45-273, 45-317, 45-425, 48-358, 49-330, 58-319, 83-327, 87-544, 87-600, 95-364, 95-652, 821-1394, 821-1421, 821-1438, 1174-1697, 1174-1880, 1274-2027, 1283-1569, 1283-1705, 1283-1904, 1292-1896, 1318-1869, 1397-1967, 1403-1904, 1432-2031, 1459-2160, 1468-2410, 1476-1904, 1494-2023, 1494-2024, 1494-2073, 1494-2076, 1494-2110, 1494-2142, 1494-2171, 1501-1904, 1503-1904, 1508-2226, 1509-1904, 1519-1904, 1520-2008, 1525-2159, 1528-2354, 1532-2277, 1542-2401, 1560-2374, 1571-1797, 1571-2357, 1583-2383, 1630-2198, 1630-2227, 1632-2201, 1654-2092, 1659-2078, 1659-2150, 1659-2159, 1659-2160, 1659-2161, 1662-2160, 1694-2527, 1720-2440, 1750-2526, 1751-2160, 1774-1952, 1783-2411, 1786-2416, 1786-2425, 1794-2527, 1810-2527, 1825-2527, 1827-2481, 1831-2160, 1837-2527, 1841-2400, 1859-2527, 1878-2527, 1883-2154, 1883-2156, 1883-2160, 1896-2527, 1908-2527, 1912-2527, 1924-2527, 1926-2527, 1936-2528, 1939-2527, 1941-2418, 1942-2761, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1950-2418, 1983-2527, 1991-2527, 2002-2527, 2043-2436, 2077-2354, 2087-2527, 2088-2287, 2088-2526, 2088-2527, 2107-2418, 2117-2527, 2128-2527, 2168-2527, 2178-2527, 2191-2407, 2208-2527, 2218-3008, 2223-2899, 2223-2968, 2223-3008, 2223-3046, 2223-3053, 2235-2419, 2236-2419, 2237-2527, 2269-2426, 2274-2526, 2279-2469, 2279-2526, 2279-2527, 2285-2419, 2286-2419, 2287-2419, 2302-2419, 2305-2594, 2370-2527, 2379-2970, 2379-3030, 2458-2909, 2458-2968, 2458-3046, 2458-3052, 2458-3061, 2458-3080, 2458-3095, 2458-3113, 2458-3170, 2556-2610, 2572-2610, 2588-2610, 2628-2968, 2629-2843, 2629-2979, 2629-3035, 2629-3144, 2651-3169, 2671-3271, 2677-3145, 2722-3419, 2813-3431, 2827-3325, 2827-3353, 2827-3385, 2827-3395, 2827-3412, 2827-3429, 2827-3448, 2827-3489, 2829-3316, 2830-3343, 2833-3407, 2833-3445, 2833-3557, 2833-3560, 2868-3562, 2880-3445, 2881-3149, 2881-3366, 2881-3428, 2881-3445, 2881-3562, 2883-3560, 2885-3437, 2887-3583, 2896-3554, 2898-3562, 2902-3562, 2903-3549, 2904-3510, 2914-3562, 2918-3562, 2923-3562, 2928-3562, 2929-3562, |
| | 2942-3333, 2942-3336, 2942-3400, 2942-3462, 2942-3474, 2942-3547, 2942-3562, 2943-3562, 2945-3242, 2945-3562, 2947-3562, 2953-3562, 2959-3562, 2965-3562, 2983-3484, 2983-3516, 2983-3558, 2983-3561, 2983-3562, 2985-3562, 2987-3562, 2990-3562, 2991-3562, 3007-3561, 3007-3562, 3030-3562, 3033-3561, 3034-3515, 3034-3534, 3034-3561, 3034-3562, 3036-3562, 3038-3562, 3044-3535, 3044-3562, 3045-3562, 3059-3562, 3060-3562, 3076-3557, 3084-3562, 3091-3561, 3096-3562, 3113-3562, 3121-3562, 3155-3562, 3273-3562, 3297-3562, 3319-3561, 3319-3562, 3320-3562, 3330-3552, 3330-3745, 3338-3561, 3338-3562, 3400-3675, 3400-3846, 3400-3873, 3400-3959, 3400-3960, 3400-4002, 3400-4060, 3400-4062, 3469-3562, 3569-3952, 3569-4044, 3569-4146, 3569-4152, 3592-4627, 3602-3791, 3639-4117, 3645-4287, 3650-4133, 3682-4066, 3712-4598, 3722-3969, 3739-4356, 3742-4357, 3751-4245, 3754-4048, 3754-4346, 3789-4460, 3824-4489, 3853-4554, 3887-4370, 3887-4384, 3888-4214, 3889-4381, 3896-4474, 3897-4223, 3904-4153, 3914-4428, 3916-4489, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 3928-4463, 3932-4607, 3935-4542, 3938-4486, 3940-4574, 3955-4192, 3959-4422, 3959-4522, 3959-4621, 3963-4514, 3964-4520, 3970-4403, 3973-4526, 3973-4603, 3997-4485, 4003-4672, 4018-4613, 4028-4296, 4028-4542, 4045-4608, 4047-4397, 4048-4566, 4049-4462, 4049-4500, 4049-4520, 4049-4560, 4049-4628, 4050-4594, 4058-4326, 4063-4310, 4064-4624, 4069-4635, 4070-4331, 4073-4570, 4079-4753, 4080-4600, 4088-4535, 4089-4822, 4095-4759, 4097-4684, 4102-4720, 4103-4366, 4116-4778, 4122-4765, 4126-4740, 4128-4338, 4131-4744, 4136-4378, 4141-4611, 4142-4450, 4154-4666, 4156-4617, 4156-4730, 4159-4347, 4161-4744, 4167-4852, 4168-4671, 4168-4799, 4171-4933, 4178-4853, 4182-4662, 4184-4853, 4189-4719, 4190-4746, 4192-4632, 4192-4721, 4195-4761, 4198-4849, 4205-4681, 4213-4473, 4213-4714, 4213-4804, 4220-4831, 4225-4892, 4228-4878, 4229-4732, 4233-4706, 4234-4705, 4235-4871; 4238-4498, 4238-4521, 4238-4719, 4238-4738, 4238-4857, 4240-4714, 4240-4822, 4241-4595, 4253-4926, 4254-4598, 4262-4769, 4266-5226, 4270-4864, 4275-4966, 4276-4619, 4276-4894, |
| | 4278-4784, 4287-4409, 4302-4808, 4304-4800, 4307-4740, 4340-4455, 4346-4958, 4349-4985, 4351-4907, 4355-4576, 4360-4566, 4371-4563, 4371-4852, 4373-4938, 4377-4707, 4381-4910, 4417-5036, 4420-4938, 4422-5035, 4424-4879, 4429-5048, 4430-5184, 4440-5059, 4445-5044, 4453-5013, 4463-4854, 4464-4715, 4468-4609, 4481-4932, 4486-5131, 4487-4659, 4489-5095, 4491-5034, 4493-5041, 4495-5141, 4509-5121, 4512-5036, 4518-4937, 4518-5147, 4527-5130, 4529-4630, 4532-5110, 4555-5125, 4559-5074, 4559-5125, 4571-5151, 4571-5192, 4572-5151, 4578-5151, 4579-5153, 4589-5279, 4599-5243, 4603-5194, 4606-4709, 4609-5168, 4613-4945, 4627-5102, 4636-5204, 4638-5058, 4638-5125, 4638-5160, 4663-5151, 4691-5001, 4710-5361, 4712-5197, 4716-4980, 4716-5190, 4717-4995, 4721-5266, 4726-5151, 4727-5360, 4738-5151, 4754-5048, 4757-5273, 4765-5151, 4767-5151, 4770-5146, 4778-5042, 4800-5219, 4823-5148, 4826-4972, 4840-5242, 4843-5146, 4858-5160, 4858-5175, 4888-5180, 4902-5241, 4906-5353, 4908-5361, 4935-5153, 4939-5362, 4956-5241, 4958-5073, 4958-5112, 4970-5361, |
| | 4990-5327, 5020-5357, 5025-5361, 5029-5267, 5029-5325, 5035-5370, 5067-5328, 5067-5333, 5067-5523, 5067-5569, 5110-5378, 5123-5273, 5153-5625, 5237-5423, 5273-5364, 5332-5584, 5334-5569, 5418-5853, 5427-5569, 5492-5569 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 77/7510082CB1/2071 | 1-704, 1-2071, 119-449, 277-950, 309-949, 401-954, 411-610, 411-616, 484-1178, 485-697, 498-1161, 563-789, 580-838, 580-1374, 581-1184, 581-1246, 689-1258, 705-1129, 707-830, 715-972, 727-924, 760-834, 762-834, 777-1035, 781-1026, 797-1098, 812-1111, 818-1018, 818-1304, 819-1102, 822-1402, 824-1100, 824-1123, 832-1243, 837-1121, 841-1260, 854-1108, 861-1159, 863-1677, 873-1142, 875-1130, 878-1131, 907-1582, 914-1183, 920-1144, 925-1234, 949-1134, 949-1156, 949-1420, 949-1436, 949-1451, 949-1469, 949-1481, 949-1486, 949-1497, 949-1521, 949-1533, 949-1590, 949-1605, 949-1613, 949-1666, 949-1729, 949-1756, 951-1571, 953-1515, 953-1599, 953-1655, 956-1200, 959-1577, 972-1247, 1001-1603, 1024-1282, 1024-1326, 1035-1334, 1052-1750, 1073-1355, 1077-1244, 1077-1302, 1079-1811, 1084-1311, 1085-1757, 1102-1373, 1102-1381, 1102-1900, 1123-1318, 1145-1344, 1145-1400, 1145-1404, 1145-1416, 1145-1820, 1146-1384, 1147-1374, 1148-1331, 1149-1442, 1151-1444, 1159-1639, 1161-2023, 1193-1444, 1205-1443, 1215-1450, 1215-1472, 1215-1727, 1219-1453, 1232-1438, |
| | 1253-1481, 1271-1852, 1276-1563, 1293-1589, 1298-1563, 1342-1984, 1351-1562, 1371-1651, 1371-1903, 1389-2064, 1398-1677, 1403-1647, 1405-1678, 1423-1900, 1443-1699, 1444-1924, 1450-1908, 1463-1722, 1475-1672, 1475-1727, 1475-1744, 1475-1766, 1479-1768, 1493-1764, 1544-1775, 1570-1713, 1570-1784 |
| 78/7510367CB1/3703 | 1-227, 1-3703, 86-856, 86-879, 2641-2911, 2671-2885, 2686-3033, 2751-3043, 2790-2999, 2790-3014, 2790-3036, 2843-3042, 2866-3497, 2902-3497, 2941-3703 |
| 79/7510413CB1/1171 | 1-1171, 229-762, 229-830, 498-987, 558-978, 579-977, 594-977, 619-977, 621-952, 646-977, 718-974, 718-987, 727-987, 753-969 |
| 80/1721303CB1/323 | 1-313, 5-290, 5-312, 35-323, 54-319, 130-323, 185-315, 186-311, 192-313, 206-312 |
| 81/7502007CB1/1221 | 1-281, 2-1086, 2-1096, 33-126, 33-212, 41-681, 41-719, 41-731, 41-758, 41-760, 41-789, 41-804, 41-821, 41-907, 41-908, 41-1048, 71-212, 71-219, 72-325, 74-212, 79-212, 93-804, 115-606, 158-1047, 189-804, 221-714, 241-1047, 246-1047, 263-374, 291-569, 390-705, 447-1047, 458-1034, 466-995, 470-1100, 478-722, 478-804, 525-1135, 547-788, 562-814, 562-816, 589-1188, 621-1101, 622-1150, 682-1067, 723-1077, 782-1221, 783-981, 788-1089, 803-1077, 803-1104, 803-1111, 914-1098 |

Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 82/7506439CB1/2008 | 1-251, 1-393, 1-2003, 94-852, 100-907, 100-920, 559-946, 559-947, 559-968, 559-991, 559-1018, 559-1098, 564-808, 565-1462, 565-1463, 570-991, 576-1462, 587-720, 604-1463, 633-1228, 635-1164, 642-1462, 676-1462, 681-1462, 714-1462, 730-1462, 732-1462, 780-1463, 816-1463, 841-1463, 857-1322, 884-1322, 897-991, 897-1466, 898-1394, 935-1213, 935-1358, 935-1390, 935-1413, 935-1441, 935-1454, 935-1471, 935-1485, 935-1504, 936-1495, 938-1470, 950-1440, 969-1164, 970-1470, 977-1196, 1034-1379, 1035-1454, 1042-1607, 1050-1576, 1067-1164, 1086-1555, 1095-1720, 1109-1676, 1124-1591, 1125-1383, 1128-1404, 1146-1904, 1146-1939, 1161-1677, 1165-1648, 1166-1677, 1167-1640, 1167-1641, 1179-1807, 1198-1607, 1219-1987, 1241-1905, 1259-1677, 1260-1673, 1281-1922, 1285-1939, 1287-1818, 1297-1769, 1303-1962, 1303-1977, 1323-1616, 1327-1971, 1334-1795, 1334-1858, 1335-1858, 1356-1873, 1361-1858, 1362-1913, 1362-1926, 1369-1633, 1396-1698, 1402-1983, 1407-1859, 1418-2008, 1424-2008, 1434-2006, 1438-1752, 1438-1792, 1455-1858, 1477-2005, 1507-2005, 1512-2008, |
| | 1516-2008, 1524-1991, 1530-2005, 1543-1991, 1554-2005, 1563-1991, 1569-2005, 1582-2005, 1585-2008, 1591-2005, 1612-1986, 1622-1991, 1625-1991, 1631-1991, 1640-2008, 1646-1991, 1656-1991, 1667-1929, 1684-1944, 1687-2002, 1692-1991, 1740-2007, 1771-2005, 1812-1980 |
| 83/7509243CB1/1080 | 1-337, 1-338, 1-1080, 19-338, 20-78, 20-133, 20-150, 20-197, 20-337, 20-338, 20-1080, 21-338, 22-338, 26-338, 29-338, 60-338, 67-338, 158-338, 171-338, 177-338, 181-338, 219-1080 |
| 84/7509404CB1/2412 | 1-236, 1-2401, 17-2402, 18-115, 18-142, 18-143, 18-178, 18-255, 18-266, 18-280, 18-281, 18-288, 19-199, 51-190, 72-204, 90-288, 127-886, 350-733, 350-780, 350-886, 350-887, 350-903, 350-922, 350-937, 350-938, 350-948, 350-949, 350-955, 350-997, 350-1010, 350-1011, 350-1038, 350-1041, 350-1044, 350-1053, 350-1058, 350-1091, 350-1119, 350-1127, 350-1163, 351-774, 415-1058, 424-699, 473-1142, 482-1183, 505-1186, 511-947, 517-795, 530-933, 588-886, 590-688, 596-1586, 617-1588, 618-1061, 646-1213, 646-1215, 651-1209, 657-877, 658-1264, 660-979, 660-1234, 674-917, 687-1297, 700-1081, 707-1267, 721-1587, 723-1274, 723-1586, 731-1291, 732-1586, 736-1211, 737-1291, 740-1254, 762-1066, 776-1587, 782-1587, 788-1588, 793-1542, 793-1586, 793-1587, 795-1167, 797-1588, 799-1586, 806-1588, 811-1112, 834-1081, 836-1588, 846-925, 846-1263, 857-1586, 866-1586, 867-1417, 869-1308, 900-1128, 922-1302, 942-1632, 945-1452, 958-1520, 988-1586, 990-1523, 1009-1622, 1024-1167, 1055-1389, 1084-1630, 1090-1361, 1117-1692, 1131-1358, 1136-1978, 1141-1757, 1146-1770, 1171-1410, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1177-1439, 1177-1442, 1201-1417, 1206-1466, 1208-1516, 1211-1931, 1225-1804, 1237-1928, 1246-1917, 1272-1862, 1276-1509, 1281-1595, 1296-1567, 1296-1687, 1307-1827, 1314-1861, 1321-1936, 1327-1610, 1346-1593, 1412-1512, 1413-1943, 1414-2231, 1417-2231, 1427-1703, 1447-1730, 1461-1869, 1468-2231, 1469-2231, 1470-1708, 1476-1949, 1477-2231, 1479-2230, 1479-2231, 1484-2228, 1515-1946, 1517-2233, 1529-1636, 1536-2231, 1547-1852, 1565-1989, 1569-1988, 1584-1989, 1585-2081, 1587-1988, 1588-1986, 1608-1987, 1616-1837, 1619-1949, 1623-1910, 1623-1988, 1656-1982, 1661-1920, 1743-1980, 1769-1982, 1775-1988, 1827-1988, 1868-2137, 1900-2231, 1910-1982, 1923-2391, 1989-2412 |
| 85/7509439CB1/1004 | 1-477, 70-479, 70-729, 77-491, 78-382, 373-1004 |
| 86/7510202CB1/5231 | 1-603, 1-747, 1-748, 1-770, 1-5226, 321-801, 394-1125, 402-982, 426-1125, 465-1126, 512-1198, 521-1118, 557-1125, 749-904, 753-1125, 1009-1419, 1009-1468, 1011-1568, 1020-1486, 1123-1569, 1123-1570, 1960-2222, 1960-2224, 2001-2382, 2002-2467, 2039-2467, 2061-2450, 2091-2454, 2151-2845, 2151-2847, 2153-2813, 2264-2917, 2329-2426, 2329-2450, 2330-2967, 2407-2781, 2407-2994, 2407-3022, 2407-3060, 2407-3100, 2407-3161, 2436-3100, 2436-3105, 2469-2987, 2636-3019, 2651-3163, 2832-3285, 2878-3258, 2958-3623, 3270-3941, 3301-3941, 3452-4125, 3577-4128, 3906-4454, 3914-4564, 4033-4638, 4143-4390, 4239-4357, 4259-4500, 4275-4523, 4342-5226, 4375-4648, 4411-4660, 4454-4890, 4719-5231 |
| 87/7510203CB1/3269 | 1-505, 1-3251, 35-625, 130-656, 132-407, 134-734, 176-653, 216-915, 217-505, 217-783, 360-416, 360-420, 360-421, 419-463, 419-479, 514-1224, 543-1147, 544-1143, 567-1078, 707-1334, 737-1224, 747-990, 747-1032, 747-1196, 747-1202, 747-1222, 862-1209, 889-1154, 932-1207, 953-1555, 968-1249, 968-1388, 968-1535, 1012-1567, 1065-1529, 1085-1256, 1132-1705, 1134-1826, 1136-1413, 1194-1842, 1195-1772, 1217-1380, 1247-1381, 1248-1773, 1267-1815, 1287-1545, 1287-1634, 1287-1792, 1287-1839, 1287-1862, 1300-1790, 1381-1943, 1381-1999, 1396-1621, 1412-1909, 1469-1686, 1472-1731, 1475-1920, 1498-2208, 1508-1704, 1527-1999, 1547-2199, 1728-2254, 1745-2291, 1756-1959, 1795-2385, 1803-2443, 1848-2429, 1851-2410, 1956-2632, 1960-2535, 1990-2656, 2009-2663, 2010-2247, 2022-2485, 2049-2543, 2058-2346, 2071-2600, 2085-2367, 2099-2778, 2104-2581, 2106-2514, 2107-2387, 2150-2395, 2159-2635, 2169-2784, 2175-2714, 2179-2549, 2185-2823, 2186-2732, 2209-2676, 2213-2867, 2246-2490, 2246-2921, 2285-2782, 2298-2824, 2298-2879, 2302-2955, 2303-2842, 2308-2600, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 2316-2427, 2330-2593, 2351-2645, 2351-2647, 2351-2659, 2360-2863, 2365-2957, 2367-2866, 2396-2845, 2404-2659, 2415-2961, 2417-3045, 2418-2858, 2421-2948, 2427-3000, 2462-3162, 2466-2983, 2512-3162, 2524-3064, 2544-3206, 2570-3193, 2578-3160, 2586-3174, 2592-3038, 2595-3236, 2603-2793, 2633-3224, 2634-3177, 2636-3093, 2645-3251, 2648-2913, 2648-3195, 2650-3224, 2650-3236, 2655-3080, 2655-3131, 2658-3245, 2671-3224, 2672-3138, 2675-3133, 2678-3261, 2688-3215, 2701-3173, 2701-3250, 2706-2966, 2706-2982, 2706-2987, 2706-2997, 2706-3000, 2706-3005, 2706-3006, 2706-3007, 2706-3011, 2706-3018, 2708-3199, 2719-3070, 2726-3224, 2727-3249, 2727-3254, 2747-3179, 2750-3016, 2750-3147, 2753-3203, 2755-3253, 2758-3224, 2766-3214, 2772-3215, 2775-3233, 2788-3045, 2791-3123, 2795-3018, 2797-3070, 2797-3254, 2814-3256, 2814-3269, 2819-3241, 2823-3241, 2829-3236, 2834-3236, 2835-3072, 2866-3236, 2868-3201, 2877-3224, 2889-3222, 2891-3224, 2903-3237, 2906-3162, 2920-3224, 2926-3238, 2976-3238, 2987-3224, 2992-3236, 2999-3231, 3014-3242, 3014-3251, |
| 88/7510208CB1/7706 | 1-635, 1-7706, 5-618, 20-353, 20-630, 22-623, 24-514, 25-545, 26-542, 26-574, 26-655, 26-659, 26-694, 29-673, 29-791, 30-139, 30-463, 30-466, 30-564, 30-566, 30-583, 30-591, 30-613, 30-623, 31-584, 31-764, 32-587, 32-655, 33-558, 35-462, 45-415, 57-690, 61-659, 62-470, 72-615, 73-230, 73-392, 75-397, 75-533, 76-567, 77-572, 77-621, 79-612, 79-647, 79-758, 80-710, 82-435, 82-628, 84-524, 84-645, 95-709, 97-678, 101-365, 171-274, 294-751, 305-670, 430-774, 492-1113, 507-987, 516-904, 536-1148, 539-1080, 582-1208, 599-1160, 660-1225, 699-998, 724-1164, 741-1305, 813-1471, 834-1478, 873-1340, 892-1482, 897-1620, 941-1452, 947-1522, 997-1375, 1058-1446, 1121-1751, 1145-1375, 1145-1452, 1145-1471, 1145-1495, 1145-1500, 1145-1521, 1145-1523, 1145-1527, 1145-1560, 1145-1598, 1145-1600, 1145-1614, 1145-1653, 1145-1661, 1145-1681, 1145-1718, 1145-1777, 1145-1830, 1203-1798, 1245-1719, 1273-1375, 1308-1507, 1376-1864, 1376-2060, 1376-2061, 1377-2061, 1379-2060, 1382-1989, 1382-2061, 1385-2063, 1391-2061, 1395-2063, 1398-2061, 1402-2061, 1403-2025, 1403-2061, 1404-2061, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1404-2063, 1408-2061, 1409-2061, 1414-2061, 1414-2063, 1417-2061, 1418-2061, 1422-2061, 1423-2061, 1426-2063, 1429-2061, 1433-2061, 1439-2061, 1442-1924, 1442-1933, 1442-1989, 1442-2030, 1442-2035, 1443-2061, 1443-2063, 1447-2061, 1447-2063, 1451-2061, 1452-2061, 1452-2063, 1453-2061, 1455-2061, 1458-2061, 1462-2061, 1467-1742, 1467-1743, 1471-2061, 1472-2061, 1474-2061, 1480-2061, 1487-2060, 1487-2061, 1487-2063, 1488-2061, 1489-2061, 1494-2061, 1496-2061, 1498-2061, 1500-2061, 1503-2060, 1504-2061, 1506-2061, 1510-2061, 1512-2061, 1515-1994, 1515-2000, 1515-2061, 1520-2061, 1524-2061, 1525-2061, 1528-2061, 1534-1992, 1541-2063, 1546-2061, 1547-2061, 1550-2061, 1557-2063, 1561-1835, 1561-1864, 1561-2061, 1564-2060, 1564-2061, 1566-2009, 1570-2061, 1572-2061, 1574-2061, 1577-2028, 1577-2061, 1580-2060, 1581-2009, 1583-2061, 1589-2066, 1596-2060, 1598-2061, 1600-2061, 1604-2060, 1610-2061, 1634-2061, 1639-2060, 1644-2060, 1644-2061, 1646-2063, 1654-2061, 1667-2060, 1684-2058, 1690-2060, 1710-2060, 1711-2060, 1736-1912, 1789-2064, |
| | 1790-2061, 1792-2059, 1850-1909, 1869-2061, 1941-2060, 1962-2061, 2068-2191, 2068-2669, 2070-2291, 2070-2606, 2071-2247, 2071-2291, 2071-2395, 2071-2451, 2071-2508, 2071-2513, 2071-2568, 2071-2576, 2071-2632, 2071-2687, 2071-2691, 2071-2715, 2071-2747, 2275-3049, 2334-2631, 2427-2543, 2662-3148, 2770-3383, 2882-3433, 2918-3447, 2979-3494, 3007-3488, 3021-3482, 3077-3576, 3078-3518, 3636-3672, 3636-3703, 3636-3704, 3643-3704, 3667-4087, 4538-5095, 4538-5381, 4580-5166, 4580-5375, 4601-5139, 4622-5204, 4638-5235, 4661-5225, 4667-5255, 4708-5336, 4711-5255, 4738-5373, 4743-5279, 4783-5295, 4784-5306, 4800-5401, 4859-5408, 4880-5279, 4884-5443, 4905-5429, 4906-5428, 4907-5392, 4911-5439, 4922-5431, 4931-5511, 4932-5539, 4943-5513, 4954-5510, 4957-5560, 4969-5513, 4990-5641, 4997-5521, 5040-5353, 5040-5427, 5041-5428, 5045-5279, 5045-5544, 5045-5772, 5054-5270, 5084-5630, 5085-5630, 5095-5545, 5117-5526, 5146-5403, 5150-5764, 5173-5458, 5174-5828, 5194-5822, 5214-5824, 5226-5872, 5234-5863, 5245-5451, 5245-5452, 5263-5869, 5267-5803, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 5269-5803, 5300-5963, 5303-5825, 5315-5482, 5315-5545, 5330-5598, 5347-5494, 5363-5985, 5376-6040, 5393-5701, 5406-5883, 5420-5784, 5420-5982, 5431-5662, 5432-5790, 5433-5678, 5449-5996, 5455-6083, 5479-6046, 5485-6220, 5507-6146, 5511-5861, 5511-6131, 5518-5800, 5554-5792, 5554-6341, 5563-5858, 5572-6335, 5582-6076, 5589-6214, 5629-5967, 5632-5813, 5665-5870, 5707-6218, 5707-6417, 5709-6233, 5731-5969, 5741-6316, 5758-6068, 5773-6039, 5773-6051, 5790-6444, 5816-6050, 5842-6403, 5846-6459, 5848-6045, 5855-6084, 5855-6091, 5855-6181, 5877-5931, 5882-6240, 5882-6381, 5885-6198, 5930-6401, 5958-6243, 5960-6226, 5968-6491, 5974-6252, 6014-6272, 6017-6502, 6032-6345, 6051-6303, 6137-6738, 6162-6360, 6163-6626, 6163-6745, 6163-6843, 6232-6660, 6248-6501, 6255-6510, 6258-6585, 6261-6536, 6277-6458, 6302-6899, 6304-6567, 6305-6784, 6307-6589, 6312-6527, 6334-6728, 6338-6595, 6338-6773, 6338-6822, 6338-6850, 6338-6892, 6338-6909, 6338-6922, 6338-6931, 6338-6990, 6338-7007, 6338-7110, 6338-7119, 6339-6937, 6339-6945, 6339-7014, 6339-7083, |
| | 6342-6898, 6342-6899, 6347-6430, 6362-6599, 6363-7043, 6366-7088, 6370-6893, 6370-6906, 6378-6847, 6387-6847, 6393-6583, 6402-6707, 6402-6847, 6402-7039, 6402-7048, 6403-6870, 6407-6671, 6412-6878, 6422-6691, 6423-6660, 6423-6673, 6424-6948; 6436-7093, 6442-7164, 6446-6908, 6446-6940, 6447-7153, 6469-6847, 6480-6729, 6484-6774, 6484-7058, 6485-6685, 6491-7000, 6492-6758, 6492-6760, 6507-7247, 6513-7239, 6515-6631, 6522-6786, 6525-7113, 6526-7057, 6549-7113, 6556-7082, 6557-6931, 6558-6831, 6558-7180, 6565-7391, 6568-7165, 6583-6691, 6587-7007, 6593-7134, 6600-6830, 6600-6845, 6600-7138, 6604-7391, 6616-7287, 6626-6855, 6626-6880, 6643-6925, 6645-6943, 6670-7202, 6678-7242, 6691-6938, 6691-7304, 6700-6982, 6702-7297, 6704-7334, 6718-6950, 6718-6951, 6729-6860, 6740-6997, 6751-7037, 6759-7327, 6764-7286, 6768-7484, 6778-7411, 6781-7335, 6781-7340, 6782-6899, 6785-7451, 6792-7055, 6820-7364, 6826-6894, 6833-7296, 6839-7239, 6839-7581, 6839-7586, 6839-7623, 6839-7663, 6839-7699, 6845-7114, 6873-7440, 6880-7256, 6882-7434, 6887-7530, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 6890-7113, 6890-7133, 6892-7394, 6898-7706, 6902-7150, 6902-7158, 6902-7333, 6910-7696, 6911-7177, 6911-7180, 6911-7509, 6920-7339, 6928-7068, 6937-7196, 6937-7516, 6942-7254, 6943-7193, 6943-7220, 6944-7483, 6944-7706, 6947-7706, 6955-7317, 6956-7508, 6977-7258, 6983-7230, 6992-7328, 6995-7251, 7004-7514, 7005-7220, 7005-7591, 7007-7694, 7007-7706, 7010-7659, 7015-7552, 7016-7706, 7022-7219, 7022-7466, 7022-7497, 7030-7272, 7030-7277, 7046-7621, 7048-7266, 7051-7343, 7051-7588, 7055-7335, 7062-7545, 7068-7349, 7068-7353, 7069-7360, 7082-7706, 7096-7629, 7098-7706, 7114-7457, 7115-7398, 7115-7706, 7119-7662, 7137-7341, 7152-7644, 7153-7664, 7156-7706, 7162-7373, 7166-7706, 7186-7655, 7210-7507, 7223-7388, 7226-7701, 7228-7515, 7230-7444, 7237-7516, 7240-7481, 7264-7584, 7265-7579, 7266-7557, 7286-7523, 7294-7560, 7302-7501, 7304-7481, 7306-7687, 7311-7562, 7311-7596, 7330-7626, 7345-7583, 7359-7602, 7385-7622, 7397-7653, 7399-7668, 7408-7654, 7409-7574, 7427-7692, 7430-7706, 7434-7676, 7435-7643, 7438-7671, 7440-7545, 7446-7683, |
| | 7450-7621, 7480-7706, 7502-7536, 7502-7537, 7502-7539, 7515-7539, 7540-7575, 7540-7577, 7541-7568, 7544-7577, 7553-757 |
| 89/7510446CB1/3159 | 1-730, 1-773, 1-3135, 1839-2568, 2165-2402, 2165-2513, 2165-2597, 2165-2628, 2165-2697, 2165-2701, 2165-2709, 2165-2725, 2165-2741, 2165-2745, 2165-2775, 2192-2671, 2196-2820, 2204-2837, 2209-2807, 2221-2959, 2236-2775, 2237-2793, 2339-3135, 2347-3002, 2380-2601, 2387-3135, 2423-2729, 2425-3135, 2426-3135, 2427-3135, 2439-2957, 2469-3135, 2475-3133, 2490-3134, 2525-3135, 2525-3154, 2550-3137, 2567-3133, 2580-3135, 2594-2811, 2615-3101, 2640-3159, 2642-2862, 2658-3133, 2665-2823, 2668-3159, 2794-2895, 2918-3089, 2992-3159 |
| 90/7505294CB1/1821 | 1-263, 1-1813, 29-558, 46-193, 110-266, 112-264, 124-485, 139-661, 254-484, 295-607, 330-646, 350-626, 365-610, 370-620, 370-848, 395-671, 426-832, 444-706, 496-704, 497-840, 507-619, 514-769, 520-800, 524-827, 566-919, 566-922, 566-1085, 581-849, 617-828, 621-870, 631-864, 641-814, 641-870, 642-858, 642-870, 643-868, 643-870, 643-928, 644-748, 644-870, 645-869, 646-870, 646-892, 648-870, 649-865, 649-870, 650-764, 651-870, 654-870, 673-870, 689-870, 691-795, 716-870, 770-870, 771-1032, 785-1060, 870-1255, 871-1439, 875-1226, 875-1257, 879-1458, 895-1100, 902-1423, 903-1140, 920-1481, 923-1257, 924-1191, 930-1407, 950-1214, 952-1215, 955-1219, 958-1225, 960-1171, 962-1226, 963-1105, 966-1206, 967-1229, 967-1252, 972-1182, 972-1428, 995-1294, 1010-1472, 1034-1202, 1041-1223, 1044-1328, 1064-1314, 1067-1311, 1071-1340, 1071-1698, 1074-1618, 1089-1328, 1089-1724, 1089-1731, 1093-1338, 1110-1373, 1110-1433, 1118-1458, 1123-1408, 1132-1325, 1132-1329, 1142-1762, 1145-1394, 1147-1668, 1148-1427, 1149-1766, 1158-1417, 1158-1745, 1160-1745, 1161-1765, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1162-1813, 1166-1440, 1169-1331, 1182-1458, 1188-1742, 1189-1413, 1191-1731, 1192-1503, 1195-1799, 1196-1478, 1200-1746, 1200-1748, 1203-1662, 1222-1497, 1233-1799, 1261-1707, 1288-1734, 1290-1390, 1316-1821, 1317-1813, 1336-1564, 1336-1622, 1344-1821, 1346-1819, 1349-1601, 1349-1718, 1353-1803, 1359-1552, 1359-1819, 1360-1633, 1360-1815, 1362-1639, 1363-1621, 1363-1768, 1366-1772, 1374-1803, 1374-1810, 1379-1761, 1389-1648, 1395-1821, 1401-1816, 1404-1698, 1406-1815, 1409-1815, 1409-1819, 1417-1819, 1427-1814, 1428-1810, 1430-1816, 1430-1819, 1433-1815, 1437-1813, 1437-1815, 1448-1695, 1450-1691, 1451-1618, 1455-1704, 1456-1815, 1463-1821, 1465-1810, 1466-1780, 1467-1817, 1471-1810, 1472-1740, 1475-1756, 1476-1817, 1478-1701, 1478-1806, 1490-1724, 1490-1766, 1493-1787, 1498-1807, 1501-1623, 1511-1810, 1512-1821, 1514-1819, 1516-1746, 1542-1807, 1542-1817, 1543-1811, 1544-1816, 1547-1799, 1548-1818, 1552-1741, 1556-1821, 1557-1759, 1559-1819, 1563-1815, 1564-1820, 1567-1805, 1568-1818, 1570-1816, 1586-1819, 1609-1820, 1612-1821, |
| | 1754-1812, 1763-1817 |
| 91/7505631CB1/3526 | 1-126, 1-242, 1-268, 1-280, 1-678, 1-3526, 3-578, 6-247, 7-501, 7-786, 14-293, 15-239, 15-246, 15-286, 15-306, 15-555, 15-556, 15-560, 18-603, 30-315, 37-643, 41-841, 95-612, 116-740, 133-811, 134-592, 160-456, 164-781, 180-840, 208-764, 283-767, 301-909, 321-883, 321-1044, 332-765, 332-968, 332-969, 403-1060, 416-976, 450-987, 450-1057, 450-1124, 450-1183, 450-1198, 450-1203, 450-1204, 450-1220, 450-1226, 450-1228, 450-1233, 450-1260, 450-1292, 450-1327, 450-1346, 450-1353, 450-1362, 459-1355, 477-877, 477-1063, 477-1165, 477-1242, 481-1071, 481-1250, 481-1263, 500-1284, 562-921, 569-1370, 594-932, 614-1079, 678-1447, 694-795, 779-1013, 779-1115, 779-1120, 779-1283, 779-1331, 779-1380, 779-1390, 779-1419, 820-1403, 852-1447, 854-1260, 876-1329, 879-1403, 933-1244, 1002-1273, 1116-1431, 1169-1446, 1197-2004, 1199-1877, 1242-1447, 1440-1657, 1440-1679, 1440-1702, 1440-1737, 1440-1946, 1440-1998, 1440-2017, 1440-2037, 1440-2077, 1463-2139, 1474-1713, 1495-1738, 1536-1830, 1543-2081, 1547-1821, 1551-1867, 1568-1798, 1572-1767, 1572-2160, 1588-1807, 1604-1835, |
| 92/7506561CB1/946 | 1-441, 1-946, 23-447, 52-261, 52-893, 176-894, 226-759, 279-759, 317-747 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 93/7510733CB1/2737 | 1-153, 1-313, 1-356, 1-384, 1-423, 1-447, 1-459, 1-490, 1-503, 1-542, 1-2737, 443-1360, 444-1360, 446-1360, 447-1360, 460-1360, 478-1360, 486-1360, 502-1361, 503-1360, 504-1360, 539-1360, 608-1100, 612-1360, 623-1360, 624-1358, 624-1360, 689-1360, 727-1124, 823-1360, 868-1088, 885-1141, 918-1143, 1015-1261, 1044-1308, 1044-1651, 1072-1676, 1107-1375, 1117-1392, 1169-1429, 1259-1760, 1282-1713, 1290-1565, 1290-1575, 1290-1776, 1290-1802, 1290-1810, 1290-1819, 1290-1871, 1290-1872, 1290-1875, 1290-1903, 1310-1947, 1319-1548, 1327-1611, 1332-1825, 1368-1615, 1369-1578, 1397-1575, 1442-1681, 1442-2013, 1456-2032, 1484-1706, 1521-1925, 1526-2057, 1567-1808, 1580-2066, 1592-1796, 1593-1924, 1599-1822, 1626-1825, 1626-1833, 1629-1867, 1648-2226, 1649-2154, 1649-2173, 1681-1917, 1681-1921, 1682-1944, 1687-1964, 1689-1896, 1689-2141, 1690-1858, 1723-2155, 1736-2226, 1747-2154, 1749-2151, 1749-2249, 1749-2304, 1750-1913, 1750-2155, 1750-2211, 1759-2464, 1760-2431, 1777-2052, 1783-2008, 1805-2027, 1805-2050, 1807-2036, 1807-2073, 1825-2093, 1827-2049, |
| | 1842-2155, 1846-2120, 1846-2155, 1847-2110, 1859-2173, 1859-2333, 1859-2356, 1859-2620, 1859-2649, 1859-2669, 1864-2672, 1868-2104, 1892-2333, 1897-2156, 1920-2162, 1944-2537, 1947-2737, 1951-2206, 1963-2457, 1980-2278, 1993-2422, 2004-2737, 2006-2737, 2018-2737, 2028-2422, 2047-2266, 2047-2287, 2047-2319, 2048-2737, 2060-2719, 2063-2294, 2072-2737, 2098-2585, 2098-2735, 2106-2657, 2108-2351, 2115-2369, 2119-2371, 2121-2548, 2122-2422, 2130-2394, 2156-2422, 2156-2636, 2156-2641, 2156-2673, 2157-2422, 2157-2673, 2172-2690, 2186-2737, 2190-2492, 2197-2443, 2207-2734, 2231-2737, 2247-2483, 2252-2732, 2253-2505, 2260-2733, 2272-2723, 2273-2730, 2273-2732, 2275-2724, 2277-2737, 2284-2556, 2284-2730, 2285-2521, 2289-2730, 2298-2731, 2313-2546, 2317-2598, 2321-2730, 2327-2732, 2333-2733, 2334-2422, 2334-2737, 2335-2737, 2336-2422, 2336-2737, 2337-2578, 2347-2730, 2347-2732, 2348-2595, 2354-2729, 2358-2578, 2361-2732, 2362-2585, 2368-2634, 2390-2733, 2403-2665, 2450-2737, 2461-2716, 2461-2729, 2476-2733, 2525-2737, 2553-2673, 2554-2696, |
| | 2560-2719, 2576-2737, 2581-2737, 2603-2729, 2675-2732, 2675-2737 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 94/7510734CB1/2821 | 1-153, 1-313, 1-356, 1-384, 1-444, 1-472, 1-2821, 66-643, 70-643, 74-628, 158-1016, 228-562, 552-1445, 702-1445, 708-1445, 709-1443, 709-1445, 774-1445, 812-1209, 908-1445, 953-1173, 970-1226, 1003-1228, 1100-1346, 1129-1393, 1129-1736, 1157-1761, 1192-1460, 1202-1477, 1254-1514, 1344-1845, 1367-1798, 1375-1650, 1375-1660, 1375-1861, 1375-1887, 1375-1895, 1375-1904, 1375-1956, 1375-1957, 1375-1960, 1375-1988, 1395-2032, 1404-1633, 1412-1696, 1417-1910, 1453-1700, 1454-1663, 1482-1660, 1527-1766, 1527-2098, 1541-2117, 1569-1791, 1606-2010, 1611-2142, 1652-1893, 1665-2151, 1677-1881, 1678-2009, 1684-1907, 1711-1910, 1711-1918, 1714-1952, 1733-2311, 1734-2239, 1734-2258, 1766-2002, 1766-2006, 1767-2029, 1772-2049, 1774-1981, 1774-2226, 1775-1943, 1808-2240, 1821-2311, 1832-2239, 1834-2236, 1834-2334, 1834-2389, 1835-1998, 1835-2240, 1835-2296, 1844-2549, 1845-2516, 1862-2137, 1868-2093, 1890-2112, 1890-2135, 1892-2121, 1892-2158, 1910-2178, 1912-2134, 1927-2240, 1931-2205, 1931-2240, 1932-2195, 1944-2258, 1944-2418, 1944-2441, 1944-2705, |
| | 1944-2734, 1944-2754, 1949-2757, 1953-2189, 1977-2418, 1982-2241, 2005-2247, 2029-2622, 2032-2821, 2036-2291, 2048-2542, 2065-2363, 2078-2507, 2089-2821, 2091-2821, 2103-2821, 2113-2507, 2132-2351, 2132-2372, 2132-2404, 2133-2821, 2145-2804, 2148-2379, 2157-2821, 2183-2670, 2183-2820, 2191-2742, 2193-2436, 2200-2454, 2204-2456, 2206-2633, 2207-2507, 2215-2479, 2241-2507, 2241-2721, 2241-2726, 2241-2758, 2242-2507, 2242-2758, 2257-2775, 2271-2821, 2275-2577, 2282-2528, 2292-2819, 2316-2821, 2332-2568, 2337-2817, 2338-2590, 2345-2818, 2357-2808, 2358-2815, 2358-2817, 2360-2809, 2362-2821, 2369-2641, 2369-2815, 2370-2606, 2374-2815, 2383-2816, 2398-2631, 2402-2683, 2406-2815, 2412-2817, 2418-2818, 2419-2507, 2419-2821, 2420-2821, 2421-2507, 2421-2821, 2422-2663, 2432-2815, 2432-2817, 2433-2680, 2439-2814, 2443-2663, 2446-2817, 2447-2670, 2453-2719, 2475-2818, 2488-2750, 2535-2821, 2546-2801, 2546-2814, 2561-2818, 2610-2821, 2638-2758, 2639-2781, 2645-2804, 2661-2821, 2666-2821, 2688-2814, 2760-2817, 2760-2821 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 95/7503977CB1/3583 | 1-674, 9-531, 28-538, 45-665, 47-3531, 94-761, 110-444, 117-452, 129-812, 137-752, 147-424, 147-432, 147-446, 147-448, 163-674, 217-929, 222-506, 232-523, 235-443, 247-929, 252-452, 268-893, 271-509, 288-833, 288-893, 288-896, 310-531, 310-820, 312-807, 326-787, 331-812, 377-920, 382-516, 474-802, 564-1199, 819-1110, 889-1134, 911-1384, 921-1079, 923-1099, 928-1517, 963-1110, 976-1651, 994-1110, 994-1490, 998-1110, 1023-1480, 1046-1658, 1077-1640, 1145-1710, 1166-1421, 1166-1527, 1166-1626, 1166-1703, 1166-1705, 1167-1505, 1184-1598, 1192-1833, 1196-1862, 1197-1704, 1210-1659, 1213-1789, 1216-1906, 1230-1900, 1241-1730, 1244-1955, 1245-1911, 1252-1910, 1254-1829, 1265-1802, 1270-1516, 1274-1922, 1275-1940, 1278-1802, 1281-1590, 1288-1923, 1292-1913, 1296-1876, 1303-1872, 1304-1868, 1306-1913, 1309-1563, 1312-1573, 1331-1854, 1337-1989, 1342-1910, 1356-1980, 1376-1998, 1382-2041, 1396-1959, 1403-2063, 1406-1787, 1421-1979, 1421-1989, 1421-2024, 1428-2101, 1433-1687, 1458-1666, 1473-1981, 1475-1730, 1479-2103, 1492-2155, 1493-2206, 1496-2190, |
| | 1510-1883, 1523-1919, 1533-2064, 1533-2185, 1535-2002, 1539-1951, 1539-2064, 1541-2138, 1542-2223, 1550-2178, 1551-2167, 1556-1930, 1576-1865, 1576-1929, 1576-2310, 1577-1930, 1577-2040, 1578-2208, 1580-2255, 1584-1955, 1584-1975, 1584-2245, 1587-2167, 1595-2242, 1598-2144, 1601-2110, 1601-2242, 1601-2348, 1602-1857, 1602-1925, 1602-1929, 1602-2086, 1602-2101, 1602-2123, 1602-2137, 1602-2221, 1602-2243, 1602-2244, 1602-2282, 1602-2289, 1602-2355, 1604-2094, 1605-2365, 1609-1865, 1615-2172, 1618-2281, 1621-2312, 1622-2169, 1622-2258, 1623-2195, 1628-2214, 1630-2342, 1638-1896, 1644-2216, 1644-2487, 1648-1974, 1649-1911, 1649-2229, 1651-2341, 1663-2113, 1664-2272, 1664-2350, 1667-1838, 1671-1838, 1678-1902, 1679-2375, 1684-1929, 1684-2120, 1684-2290, 1684-2305, 1684-2348, 1684-2395, 1685-2143, 1686-1935, 1688-2312, 1689-2429, 1692-2366, 1694-1923, 1694-2206, 1710-2217, 1722-2075, 1727-1942, 1734-2104, 1737-2330, 1739-2255, 1747-2379, 1748-2338, 1753-2447, 1754-1904, 1755-2391, 1756-2423, 1757-2072, 1757-2453, 1760-2142, 1763-2424, |

Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1767-2001, 1767-2442, 1770-2316, 1772-2462, 1775-2453, 1781-2430, 1782-2363, 1790-1854, 1796-2105, 1817-2500, 1823-2342, 1831-2110, 1836-2501, 1837-2496, 1837-2529, 1856-2545, 1865-2005, 1876-2415, 1881-2322, 1888-2106, 1888-2120, 1893-2562, 1898-2218, 1902-2334, 1919-2535, 1920-2636, 1923-2500, 1924-2266, 1933-2141, 1936-2146, 1943-2601, 1945-2064, 1948-2076, 1954-2587, 1956-2558, 1963-2115, 1965-2425, 1970-2648, 1971-2299, 1982-2646, 1989-2684, 1995-2631, 2000-2334, 2003-2550, 2003-2652, 2004-2455, 2004-2535, 2004-2655, 2006-2624, 2007-2358, 2011-2415, 2012-2658, 2014-2249, 2014-2730, 2019-2119, 2019-2625, 2022-2355, 2025-2695, 2027-2497, 2029-2323, 2029-2662, 2031-2669, 2038-2703, 2039-2557, 2042-2529, 2047-2714, 2051-2678, 2055-2332, 2065-2593, 2068-2655, 2069-2296, 2073-2338, 2080-2408, 2082-2352, 2082-2380, 2082-2529, 2082-2544, 2082-2716, 2082-2847, 2085-2581, 2107-2335, 2110-2334, 2138-2397, 2139-2720, 2141-2360, 2146-2402, 2155-2634, 2164-2798, 2171-2519, 2175-2735, 2182-3014, 2196-2455, 2208-2662, 2225-2879, 2226-2699, |
| | 2226-2826, 2226-2845, 2235-2867, 2241-2718, 2252-2695, 2263-2601, 2274-2655, 2275-2507, 2276-2867, 2282-2936, 2285-2915, 2296-2509, 2297-2580, 2304-2948, 2309-2867, 2309-2997, 2310-2844, 2310-2956, 2312-2525, 2312-2834, 2321-2418, 2331-2866, 2335-2540, 2339-2635, 2341-2498, 2343-2590, 2346-2936, 2362-2601, 2364-3036, 2366-2601, 2366-2646, 2366-2649, 2370-2657, 2371-2575, 2371-2597, 2372-3176, 2374-2624, 2374-2646, 2374-2658, 2374-2672, 2376-2588, 2376-2644, 2376-2651, 2377-2610, 2378-2633, 2378-2696, 2378-2729, 2379-2609, 2379-3044, 2381-2649, 2381-2650, 2386-2589, 2390-2673, 2396-2679, 2396-2686, 2398-2672, 2398-3084, 2418-3064, 2421-2662, 2421-2781, 2440-3147, 2445-2654, 2445-2665, 2448-2724, 2452-2985, 2454-3000, 2455-2688, 2458-2691, 2458-2739, 2459-2703, 2459-3201, 2462-3050, 2468-3031, 2472-2688, 2479-2985, 2482-2737, 2482-3040, 2483-2744, 2484-2737, 2494-2749, 2495-2607, 2496-3092, 2507-2796, 2522-2745, 2522-2748, 2527-2781, 2529-3090, 2542-3085, 2576-2815, 2582-3134, 2586-3108, 2623-2829, 2629-3142, 2635-3253, 2640-2906, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 2654-3369, 2660-2917, 2670-3209, 2672-2990, 2672-3004, 2692-2906, 2699-2946, 2700-3042, 2710-2951, 2710-3281, 2716-2980, 2716-3014, 2716-3327, 2716-3339, 2717-2957, 2718-2964, 2718-3004, 2718-3316, 2718-3322, 2720-3337, 2723-2806, 2730-2993, 2730-3056, 2731-2975, 2737-3019, 2738-2978, 2756-3067, 2765-3074, 2765-3084, 2768-3022, 2769-3538, 2785-3467, 2786-3042, 2790-3551, 2811-3085, 2819-3121, 2819-3269, 2838-3137, 2845-3432, 2855-3081, 2863-3553, 2865-3096, 2875-3180, 2880-3173, 2887-3425, 2890-3552, 2893-3176, 2894-3524, 2900-3327, 2905-3552, 2907-3530, 2911-3170, 2911-3424, 2916-3465, 2920-3395, 2923-3553, 2924-3496, 2932-3553, 2936-3006, 2936-3212, 2938-3194, 2942-3523, 2944-3231, 2944-3552, 2945-3382, 2959-3198, 2960-3499, 2964-3196, 2970-3550, 2972-3153, 2985-3235, 2985-3252, 2990-3552, 2992-3243, 2997-3551, 2998-3242, 3003-3276, 3008-3554, 3016-3519, 3028-3550, 3032-3323, 3035-3496, 3054-3313, 3054-3328, 3060-3545, 3066-3479, 3066-3480, 3066-3553, 3071-3328, 3072-3541, 3073-3553, 3075-3553, 3082-3551, 3084-3245, 3085-3300, |
| | 3085-3339, 3088-3552, 3093-3538, 3096-3312, 3111-3543, 3112-3541, 3113-3552, 3113-3554, 3114-3530, 3126-3399, 3129-3583, 3154-3542, 3156-3540, 3166-3535, 3167-3468, 3188-3438, 3192-3338, 3196-3434, 3204-3558, 3223-3436, 3224-3504, 3226-3537, 3233-3541, 3234-3506, 3235-3539, 3239-3523, 3243-3510, 3260-3482, 3260-3492, 3265-3549, 3268-3538, 3276-3535, 3279-3537, 3298-3524, 3298-3538, 3325-3538, 3327-3517, 3370-3582 |
| 96/7505084CB1/2125 | 1-210, 1-372, 1-383, 1-412, 1-418, 1-456, 1-479, 1-499, 1-500, 1-2125, 18-611, 408-872, 551-1187, 567-1115, 603-1138, 636-916, 636-1077, 636-1107, 636-1124, 636-1210, 636-1288, 639-1288, 688-1149, 745-1352, 878-1429, 1060-1230, 1060-1249, 1104-1966, 1184-1678, 1184-1689, 1184-1738, 1228-1825, 1230-1595, 1241-1858, 1294-1664, 1298-1799, 1302-1920, 1388-2125, 1419-1728, 1420-1936, 1449-2125, 1463-2111, 1470-1920, 1487-1739, 1496-2076, 1547-2116, 1550-2015, 1559-1719, 1559-1849, 1579-2125, 1664-2121, 1715-1949, 1732-2103, 1732-2121, 1738-1997, 1780-1933, 1792-2122, 1798-2121, 1836-2086, 1888-2125, 1929-2120 |
| 97/7506950CB1/1517 | 1-479, 1-1517, 37-229, 37-849, 98-1037, 240-490, 507-881, 518-1033, 519-1033, 542-1037, 796-1285, 838-1248, 838-1331, 840-1095, 1053-1403 |
| 98/7506951CB1/1694 | 1-479, 1-1694, 37-229, 37-596, 37-603, 37-657, 37-676, 37-698, 37-699, 37-701, 37-711, 37-723, 37-730, 37-740, 37-741, 37-747, 37-754, 37-755, 37-765, 37-776, 37-778, 37-782, 37-819, 37-847, 37-865, 37-877, 37-880, 37-907, 38-728, 39-744, 40-662, 240-490, 303-766, 427-1214, 436-1213, 491-1214, 568-1214, 973-1462, 1015-1425, 1015-1508, 1017-1272, 1230-1580 |
| 99/7506954CB1/1102 | 1-1102, 114-310, 114-622, 381-870, 423-833, 423-916, 425-680, 638-988 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 100/7506956CB1/1744 | 1-479, 1-1744, 37-231, 37-270, 37-596, 37-603, 37-657, 37-676, 37-698, 37-699, 37-701, 37-723, 37-730, 37-740, 37-741, 37-747, 37-763, 37-766, 37-903, 38-728, 39-744, 40-662, 240-490, 303-747, 331-1264, 434-1264, 553-1264, 622-644, 743-1260, 746-1260, 769-1264, 1023-1512, 1065-1475, 1065-1558, 1067-1322, 1280-1630 |
| 101/7506959CB1/1753 | 1-479, 1-1753, 37-231, 37-270, 37-596, 37-913, 240-490, 476-1273, 598-1269, 603-974, 603-1269, 603-1273, 607-1273, 648-1213, 651-1273, 682-1273, 754-1269, 755-1269, 778-1273, 1032-1521, 1074-1484, 1074-1567, 1076-1331, 1289-1639 |
| 102/7506960CB1/1609 | 1-1609, 114-323, 114-744, 114-914, 114-926, 114-942, 114-981, 114-982, 206-1125, 307-1129, 368-1129, 394-1129, 397-1129, 612-1125, 634-1129, 888-1377, 930-1340, 930-1423, 932-1187, 1145-1495 |
| 103/7510540CB1/1930 | 1-133, 1-134, 1-742, 1-874, 1-1930, 229-892, 328-772, 440-835, 454-720, 465-1137, 524-761, 541-960, 567-752, 613-1194, 616-1218, 623-974, 649-770, 656-981, 684-1488, 735-956, 790-1605, 804-963, 807-963, 815-1607, 861-1607, 950-1345, 975-1607, 1028-1449, 1179-1453, 1234-1511, 1363-1647, 1363-1893, 1384-1652, 1414-1814, 1441-1704, 1482-1815, 1515-1786, 1646-1899 |
| 104/7510545CB1/1205 | 1-789, 4-1205, 429-640, 429-644, 429-650, 429-654, 429-661, 429-666, 429-669, 429-704, 429-799, 429-900, 429-989, 429-1060, 429-1084, 435-716, 435-1036, 436-818, 450-605, 454-992, 454-1150, 456-706, 456-710, 471-996, 509-1083, 519-795, 532-1156, 539-1056, 553-751, 553-796, 553-805, 556-811, 556-1064, 556-1135, 558-822, 559-834, 564-1069, 569-853, 570-1012, 572-849, 582-855, 583-802, 583-1107, 588-858, 598-1071, 599-831, 604-904, 610-855, 627-913, 630-873, 631-839, 634-859, 635-832, 635-905, 640-804, 644-948, 652-850, 672-892, 676-961, 678-952, 682-1038, 687-848, 691-887, 691-957, 691-980, 704-985, 707-980, 707-985, 713-951, 719-948, 730-964, 732-914, 732-1147, 735-991, 735-1018, 748-1003, 749-1143, 750-988, 757-1142, 766-1021, 777-1043, 777-1080, 779-1030, 780-1052, 783-1036, 799-1030, 799-1040, 800-1046, 801-934, 803-1067, 807-1155, 813-1071, 815-1082, 815-1104, 823-1113, 827-1018, 833-1114, 838-1107, 838-1111, 843-1110, 843-1112, 851-1088, 851-1089, 851-1116, 859-1056, 859-1151, 859-1155, 865-1054, 866-1114, 871-1155, 883-1155, 887-1057, 940-1113, |
| | 950-1071 |

Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 105/7510654CB1/1790 | 1-260, 2-313, 2-1789, 20-595, 20-732, 20-772, 40-447, 41-763, 41-838, 41-942, 43-788, 43-811, 43-900, 43-944, 111-335, 111-449, 121-258, 129-501, 129-592, 156-610, 222-595, 222-716, 222-983, 279-860, 360-593, 360-893, 420-747, 452-753, 455-575, 524-919, 529-941, 530-1104, 544-772, 583-832, 645-848, 663-935, 663-1023, 675-990, 703-966, 711-973, 720-1023, 731-946, 766-999, 776-914, 831-1023, 843-871, 919-1023, 1022-1330, 1022-1544, 1027-1579, 1029-1287, 1029-1492, 1029-1519, 1029-1566, 1029-1570, 1029-1598, 1034-1277, 1040-1578, 1041-1317, 1045-1579, 1048-1578, 1055-1579, 1060-1266, 1060-1337, 1063-1256, 1063-1309, 1065-1434, 1065-1579, 1069-1372, 1073-1214, 1073-1286, 1073-1300, 1074-1533, 1076-1311, 1079-1578, 1081-1579, 1082-1579, 1089-1574, 1090-1347, 1093-1291, 1093-1729, 1097-1579, 1105-1573, 1108-1557, 1108-1576, 1116-1339, 1125-1391, 1130-1408, 1132-1579, 1138-1308, 1157-1385, 1161-1401, 1161-1579, 1173-1404, 1196-1394, 1204-1576, 1211-1454, 1221-1576, 1228-1345, 1243-1519, 1244-1655, 1259-1529, 1259-1710, 1272-1789, 1277-1578, 1278-1579, |
| | 1381-1789, 1395-1579, 1396-1578, 1397-1579, 1422-1576, 1451-1579, 1480-1579, 1712-1786 |
| 106/7510660CB1/3824 | 1-3820, 119-822, 124-917, 140-548, 140-813, 149-617, 166-895, 168-718, 501-1113, 705-1045, 751-1354, 758-994, 768-1517, 787-1528, 816-1535, 850-1078, 850-1138, 850-1303, 850-1436, 850-1489, 850-1491, 850-1541, 853-1499, 883-1266, 883-1406, 883-1423, 883-1467, 883-1470, 903-1470, 909-1444, 971-1692, 991-1664, 1010-1262, 1057-1604, 1070-1814, 1104-1465, 1104-1470, 1105-1470, 1115-1470, 1126-1470, 1164-1465, 1174-1470, 1178-1465, 1190-1470, 1199-1461, 1199-1470, 1214-1470, 1218-1470, 1220-1470, 1220-1841, 1226-1470, 1239-1465, 1284-1812, 1303-1799, 1333-1891, 1340-1958, 1368-1986, 1376-1465, 1398-1951, 1399-1780, 1399-1873, 1399-1913, 1399-1944, 1399-2015, 1401-1998, 1401-2002, 1403-2175, 1404-1576, 1415-2104, 1418-2059, 1489-2107, 1497-1941, 1506-2095, 1525-1769, 1525-2028, 1525-2054, 1529-1622, 1529-1768, 1529-1775, 1529-2124, 1529-2135, 1536-2082, 1543-1794, 1543-1827, 1546-1711, 1548-1831, 1563-2137, 1576-2159, 1590-1850, 1629-1941, 1653-2228, 1657-1823, 1687-1930, 1687-2150, 1687-2173, 1687-2229, 1699-2228, 1703-1848, 1727-2001, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1748-2229, 1753-1941, 1787-2192, 1788-2036, 1824-2105, 1851-1941, 1967-2212, 1974-2229, 1985-2169, 2001-2207, 2010-2317, 2010-2353, 2010-2403, 2010-2614, 2010-2819, 2096-2858, 2106-2621, 2116-2718, 2168-2821, 2172-2859, 2193-2801, 2214-2888, 2228-2484, 2228-2496, 2229-2825, 2233-2726, 2234-2695, 2236-2854, 2237-2904, 2239-2584, 2243-2569, 2249-2501, 2262-2797, 2262-2853, 2271-2768, 2272-2601, 2274-2756, 2274-2785, 2278-2808, 2280-2623, 2288-2448, 2294-2876, 2296-2887, 2301-2895, 2302-2695, 2308-2588, 2321-2982, 2352-2628, 2360-3029, 2364-3005, 2367-2948, 2367-3028, 2368-2671, 2421-2971, 2421-2985, 2424-2852, 2425-2597, 2427-3098, 2439-3064, 2446-2710, 2452-2734, 2452-2769, 2453-2921, 2459-2780, 2461-3130, 2462-3135, 2467-2672, 2468-3065, 2473-3067, 2481-2883, 2482-2883, 2487-2862, 2488-3089, 2495-3097, 2497-3221, 2502-2765, 2510-3255, 2518-2998, 2520-3207, 2521-3017, 2522-2769, 2536-2807, 2536-3026, 2537-3053, 2541-3046, 2541-3052, 2541-3080, 2542-3320, 2543-2810, 2544-2784, 2545-3156, 2546-2839, 2551-3164, 2555-3137, 2557-3065, |
| | 2557-3068, 2563-2746, 2566-3245, 2570-3102, 2575-2971, 2580-3183, 2589-3050, 2589-3296, 2590-3277, 2593-3296, 2596-2888, 2611-2949, 2612-2871, 2615-3184, 2616-2918, 2617-3282, 2620-2809, 2620-3195, 2626-3136, 2636-2925, 2649-3124, 2654-3130, 2658-3168, 2666-2991, 2669-3137, 2672-2902, 2675-2933, 2678-2951, 2684-3175, 2689-3342, 2690-2961, 2699-2968, 2701-3310, 2721-3355, 2722-3151, 2723-3199, 2725-3282, 2727-3116, 2728-3233, 2729-3116, 2734-3212, 2737-2999, 2739-3233, 2755-3395, 2756-3330, 2758-3018, 2760-3056, 2761-2931, 2764-2901, 2770-3076, 2794-3169, 2794-3285, 2794-3401, 2795-3047, 2795-3050, 2795-3086, 2811-3303, 2812-3368, 2813-3679, 2820-3304, 2826-3397, 2827-3339, 2833-3067, 2841-3091, 2841-3244, 2841-3307, 2847-3408, 2855-3267, 2885-3125, 2894-3134, 2894-3141, 2912-3161, 2912-3363, 2918-3263, 2918-3458, 2920-3291, 2927-3242, 2949-3360, 2949-3431, 2953-3425, 2958-3293, 2958-3506, 2967-3206, 2972-3259, 2975-3289, 2975-3294, 2986-3102, 2988-3609, 2988-3686, 3009-3288, 3024-3250, 3029-3815, 3038-3506, 3059-3318, 3059-3649, |

Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 3059-3660, 3061-3298, 3062-3275, 3064-3327, 3071-3311, 3071-3615, 3078-3361, 3082-3574, 3086-3374, 3089-3329, 3089-3355, 3095-3506, 3102-3506, 3116-3643, 3125-3689, 3130-3306, 3132-3815, 3135-3578, 3139-3419, 3141-3815, 3145-3764, 3149-3459, 3153-3391, 3160-3807, 3161-3466, 3162-3413, 3164-3815, 3165-3374, 3165-3415, 3165-3509, 3168-3824, 3170-3460, 3172-3278, 3173-3442, 3174-3424, 3174-3824, 3181-3401, 3184-3792, 3185-3474, 3187-3394, 3197-3461, 3197-3719, 3212-3411, 3212-3459, 3225-3815, 3227-3387, 3228-3685, 3229-3469, 3230-3535, 3240-3478, 3245-3743, 3245-3777, 3247-3773, 3250-3779, 3277-3794, 3287-3742, 3287-3824, 3296-3824, 3300-3685, 3303-3815, 3308-3778, 3312-3824, 3315-3449, 3323-3814, 3325-3506, 3326-3506, 3329-3824, 3331-3506, 3340-3506, 3344-3824, 3357-3572, 3360-3824, 3367-3824, 3369-3815, 3371-3620, 3372-3824, 3377-3506, 3377-3598, 3377-3667, 3377-3763, 3377-3798, 3377-3815, 3386-3821, 3388-3629, 3388-3816, 3392-3646, 3392-3815, 3402-3652, 3407-3815, 3411-3775, 3411-3815, 3417-3606, 3418-3817, 3420-3659, 3420-3824, |
| | 3426-3822, 3426-3824, 3429-3816, 3430-3816, 3437-3505, 3438-3503, 3441-3506, 3441-3687, 3441-3702, 3441-3704, 3441-3705, 3441-3706, 3441-3733, 3442-3506, 3444-3506, 3446-3506, 3457-3506, 3461-3506, 3463-3506, 3465-3506, 3466-3506, 3468-3506, 3470-3506, 3473-3506, 3474-3755, 3475-3506, 3478-3506, 3486-3506, 3488-3510, 3505-3525, 3505-3544, 3505-3546, 3505-3554, 3505-3565, 3505-3566, 3505-3567, 3505-3658, 3505-3684, 3505-3685, 3505-3689, 3505-3696, 3505-3699, 3505-3702, 3505-3704, 3505-3707, 3505-3716, 3505-3718, 3505-3720, 3505-3724, 3505-3734, 3505-3735, 3505-3752, 3505-3778, 3505-3800, 3505-3811, 3505-3815, 3505-3818, 3505-3819, 3505-3820, 3505-3821, 3505-3822, 3505-3824, 3506-3816, 3507-3769, 3507-3801, 3524-3820, 3531-3816, 3536-3815, 3539-3777, 3549-3796, 3554-3776, 3561-3815, 3566-3784, 3566-3820, 3573-3775, 3589-3815, 3598-3816, 3620-3815, 3647-3816, 3648-3818, 3649-3776, 3649-3815, 3651-3815, 3674-3815, 3674-3821, 3705-3824, 3708-3824, 3736-3819 |

Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 107/7510661CB1/3770 | 1-3770, 119-822, 124-917, 140-548, 140-813, 149-617, 166-895, 168-718, 501-1113, 705-1045, 751-1354, 758-994, 768-1517, 787-1528, 816-1535, 850-1078, 850-1138, 850-1303, 850-1436, 850-1489, 850-1491, 850-1541, 853-1499, 883-1266, 883-1406, 883-1423, 883-1467, 883-1470, 903-1470, 909-1444, 971-1692, 991-1664, 1010-1262, 1057-1604, 1070-1814, 1104-1465, 1104-1470, 1105-1470, 1115-1470, 1126-1470, 1164-1465, 1174-1470, 1178-1465, 1190-1470, 1199-1461, 1199-1470, 1214-1470, 1218-1470, 1220-1470, 1220-1841, 1226-1470, 1239-1465, 1284-1812, 1303-1799, 1333-1891, 1340-1941, 1354-2064, 1368-1997, 1376-1465, 1399-1780, 1399-1873, 1399-1913, 1399-1941, 1404-1576, 1445-2098, 1489-2247, 1525-1769, 1529-1622, 1529-1768, 1529-1775, 1543-1794, 1543-1827, 1546-1711, 1548-1831, 1590-1850, 1629-2283, 1657-1823, 1686-2379, 1687-1930, 1703-1848, 1753-2333, 1843-2476, 1849-2545, 1851-2544, 1886-2540, 1914-2573, 1939-2749, 1940-2099, 1940-2137, 2026-2788, 2036-2551, 2046-2648, 2098-2751, 2102-2789, 2123-2731, 2144-2818, 2158-2414, 2158-2426, 2159-2755, |
| | 2163-2656, 2164-2625, 2166-2784, 2167-2834, 2169-2514, 2173-2499, 2179-2431, 2192-2727, 2192-2783, 2201-2698, 2202-2531, 2204-2686, 2204-2715, 2208-2738, 2210-2553, 2218-2378, 2224-2806, 2226-2817, 2231-2825, 2232-2625, 2238-2518, 2251-2912, 2282-2558, 2290-2959, 2294-2935, 2297-2878, 2297-2958, 2298-2601, 2351-2901, 2351-2915, 2354-2782, 2355-2527, 2357-3028, 2369-2994, 2376-2640, 2382-2664, 2382-2699, 2383-2851, 2389-2710, 2391-3060, 2392-3065, 2397-2602, 2398-2995, 2403-2997, 2411-2813, 2412-2813, 2417-2792, 2418-3019, 2425-3027, 2427-3151, 2432-2695, 2440-3185, 2448-2928, 2450-3137, 2451-2947, 2452-2699, 2466-2737, 2466-2956, 2467-2983, 2471-2976, 2471-2982, 2471-3010, 2472-3250, 2473-2740, 2474-2714, 2475-3086, 2476-2769, 2481-3094, 2485-3067, 2487-2995, 2487-2998, 2493-2676, 2496-3175, 2500-3032, 2505-2901, 2510-3113, 2519-2980, 2519-3226, 2520-3207, 2523-3226, 2526-2818, 2541-2879, 2542-2801, 2545-3114, 2546-2848, 2547-3212, 2550-2739, 2550-3125, 2556-3066, 2566-2855, 2579-3054, 2584-3060, 2588-3098, 2596-2921, 2599-3067, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 2602-2832, 2605-2863, 2608-2881, 2614-3105, 2619-3272, 2620-2891, 2629-2898, 2631-3240, 2651-3285, 2652-3081, 2653-3129, 2655-3212, 2657-3046, 2658-3163, 2659-3046, 2664-3142, 2667-2929, 2669-3163, 2685-3325, 2686-3260, 2688-2948, 2690-2986, 2691-2861, 2694-2831, 2700-3006, 2724-3099, 2724-3215, 2724-3331, 2725-2977, 2725-2980, 2725-3016, 2741-3233, 2742-3298, 2743-3630, 2750-3234, 2756-3327, 2757-3269, 2763-2997, 2771-3021, 2771-3174, 2771-3237, 2777-3338, 2785-3197, 2815-3055, 2824-3064, 2824-3071, 2842-3091, 2842-3293, 2848-3193, 2848-3388, 2850-3221, 2857-3172, 2879-3290, 2879-3361, 2883-3355, 2888-3223, 2888-3437, 2897-3136, 2902-3189, 2905-3219, 2905-3224, 2916-3032, 2918-3560, 2918-3637, 2939-3218, 2954-3180, 2959-3766, 2968-3520, 2989-3248, 2989-3600, 2989-3611, 2991-3228, 2992-3205, 2994-3257, 3001-3241, 3001-3566, 3008-3291, 3012-3525, 3016-3304, 3019-3259, 3019-3285, 3025-3453, 3032-3518, 3046-3594, 3055-3640, 3062-3766, 3069-3349, 3071-3766, 3075-3715, 3079-3389, 3083-3321, 3090-3758, 3091-3396, 3092-3343, 3094-3766, |
| | 3095-3304, 3095-3345, 3095-3440, 3098-3770, 3100-3390, 3102-3208, 3103-3372, 3104-3354, 3104-3770, 3111-3331, 3114-3743, 3115-3404, 3117-3324, 3127-3391, 3127-3670, 3142-3341, 3142-3389, 3155-3766, 3157-3317, 3158-3636, 3159-3399, 3160-3432, 3170-3408, 3175-3694, 3175-3728, 3177-3724, 3180-3730, 3207-3745, 3217-3693, 3217-3767, 3226-3770, 3230-3636, 3233-3766, 3238-3729, 3242-3770, 3253-3765, 3255-3495, 3256-3497, 3259-3770, 3261-3516, 3270-3505, 3271-3402, 3274-3770, 3287-3523, 3290-3770, 3297-3770, 3299-3766, 3301-3571, 3302-3770, 3307-3476, 3307-3618, 3307-3714, 3307-3749, 3307-3766, 3316-3770, 3318-3580, 3318-3767, 3322-3597, 3322-3766, 3332-3603, 3337-3766, 3341-3726, 3347-3557, 3348-3768, 3350-3610, 3350-3770, 3356-3768, 3356-3770, 3359-3767, 3360-3767, 3367-3770, 3368-3770, 3371-3638, 3371-3647, 3371-3653, 3371-3655, 3371-3656, 3371-3657, 3371-3684, 3372-3655, 3372-3658, 3372-3685, 3372-3703, 3374-3653, 3376-3729, 3387-3762, 3391-3766, 3393-3770, 3395-3766, 3396-3769, 3398-3669, 3400-3609, 3400-3636, 3400-3650, 3403-3766, |
| | 3403-3770, 3405-3667, 3405-3671, 3408-3517, 3410-3770, 3416-3640, 3418-3769, 3428-3751, 3431-3635, 3433-3770, 3435-3766, 3435-3769, 3437-3766, 3438-3770, 3440-3766, 3440-3770, 3442-3686, 3445-3770, 3446-3766, 3447-3766, 3447-3770, 3449-3675, 3452-3770, 3457-3767, 3458-3720, 3458-3752, 3475-3770, 3482-3767, 3487-3766, 3490-3728, 3500-3747, 3505-3727, 3512-3766, 3517-3735, 3517-3770, 3524-3726, 3540-3766, 3549-3767, 3571-3766, 3598-3767, 3599-3769, 3600-3727, 3600-3766, 3602-3766, 3625-3766, 3625-3770, 3656-3770, 3659-3770, 3687-3770 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 108/7510680CB1/1978 | 1-350, 12-308, 12-1976, 131-408, 140-507, 521-651, 521-652, 521-782, 521-892, 521-972, 521-982, 521-1005, 521-1086, 594-866, 598-652, 618-1152, 654-1072, 703-1311, 721-1280, 740-929, 755-1354, 771-1310, 786-1382, 792-1096, 798-1267, 801-1389, 807-1294, 809-1078, 809-1452, 822-1497, 858-1322, 859-1246, 864-1427, 871-1349, 873-1155, 873-1159, 882-1472, 893-1241, 896-1443, 902-1315, 937-1513, 1028-1427, 1034-1330, 1061-1310, 1118-1375, 1298-1513, 1335-1830, 1510-1856, 1510-1870, 1510-1883, 1510-1972, 1510-1975, 1510-1976, 1510-1978, 1534-1962, 1560-1949, 1618-1961 |
| 109/7505145CB1/1622 | 1-153, 1-236, 1-1612, 7-243, 9-238, 9-491, 9-570, 14-257, 14-289, 14-320, 21-228, 21-268, 25-284, 27-228, 27-238, 27-279, 30-319, 32-265, 32-272, 32-288, 33-275, 33-290, 36-328, 39-219, 39-232, 39-369, 40-307, 40-501, 41-264, 41-324, 41-331, 41-348, 42-308, 42-310, 44-312, 44-548, 44-626, 45-229, 45-282, 47-335, 51-199, 51-307, 56-312, 60-303, 71-683, 74-654, 76-331, 83-331, 83-339, 90-303, 91-242, 94-321, 99-326, 106-375, 126-427, 151-640, 157-887, 159-435, 177-627, 189-447, 198-480, 199-442, 272-521, 304-881, 326-533, 354-577, 357-603, 357-859, 357-902, 359-774, 360-522, 361-630, 363-884, 377-638, 390-645, 395-637, 398-676, 408-579, 427-648, 429-614, 430-923, 430-994, 443-701, 451-639, 456-699, 456-715, 461-679, 465-719, 477-740, 477-755, 487-681, 493-954, 493-965, 512-777, 528-794, 532-743, 538-1070, 548-815, 553-691, 560-879, 577-804, 583-963, 615-839, 618-861, 624-835, 643-809, 644-872, 655-971, 671-933, 694-979, 699-994, 702-970, 702-994, 728-994, 746-976, 754-977, 764-1019, 767-994, 773-862, 777-915, 783-992, 797-971, 811-954, 949-1603, 994-1218, 994-1240, 994-1274, |
| | 994-1278, 994-1331, 994-1568, 994-1616, 995-1614, 996-1256, 996-1564, 997-1212, 997-1274, 1004-1318, 1005-1261, 1005-1519, 1008-1555, 1013-1612, 1014-1318, 1025-1622, 1037-1294, 1037-1622, 1058-1312, 1062-1248, 1068-1621, 1071-1622, 1074-1622, 1087-1379, 1094-1325, 1097-1350, 1103-1594, 1105-1615, 1114-1333, 1114-1418, 1116-1251, 1116-1569, 1117-1360, 1123-1360, 1124-1389, 1133-1436, 1135-1611, 1142-1622, 1143-1440, 1153-1622, 1162-1387, 1164-1614, 1171-1622, 1177-1622, 1182-1622, 1192-1611, 1194-1614, 1195-1459, 1195-1467, 1196-1622, 1198-1526, 1199-1610, 1200-1609, 1201-1610, 1206-1614, 1209-1610, 1212-1610, 1214-1610, 1215-1609, 1218-1610, 1220-1612, 1221-1610, 1224-1600, 1230-1610, 1232-1375, 1235-1609, 1235-1612, 1236-1415, 1238-1609, 1239-1622, 1240-1610, 1246-1622, 1248-1589, 1258-1510, 1268-1614, 1270-1517, 1270-1609, 1270-1622, 1273-1599, 1277-1611, 1282-1609, 1283-1620, 1284-1613, 1284-1622, 1292-1542, 1295-1561, 1300-1571, 1301-1609, 1302-1438, 1304-1619, 1309-1609, 1313-1564, 1319-1581, 1330-1578, 1337-1616, 1346-1607, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1350-1582, 1353-1611, 1363-1610, 1371-1611, 1371-1622, 1399-1608, 1399-1609, 1407-1622, 1412-1610, 1415-1619, 1415-1622, 1416-1609, 1417-1599, 1418-1615, 1419-1614, 1421-1609, 1422-1611, 1422-1612, 1424-1611, 1431-1576, 1436-1603, 1444-1619, 1444-1622, 1453-1583, 1453-1596, 1454-1622, 1476-1611, 1479-1615, 1496-1613, 1521-1615, 1522-1622 |
| 110/7505162CB1/1982 | 1-256, 1-294, 1-439, 1-472, 1-486, 1-490, 1-533, 1-545, 1-663, 2-275, 4-775, 5-212, 5-292, 5-1982, 7-245, 7-253, 7-279, 7-280, 7-339, 9-264, 10-689, 14-308, 18-299, 19-258, 33-287, 60-338, 89-327, 106-409, 113-413, 113-495, 113-672, 114-371, 116-350, 139-671, 153-399, 157-668, 158-422, 159-562, 173-342, 180-342, 189-720, 192-467, 229-870, 263-755, 285-628, 289-547, 315-873, 333-592, 333-692, 335-602, 342-579, 354-858, 369-590, 369-839, 375-983, 382-652, 401-631, 401-999, 443-1052, 446-958, 505-754, 529-999, 594-748, 594-856, 602-781, 602-804, 605-919, 609-827, 609-830, 637-910, 640-972, 671-999, 697-949, 704-961, 788-1349, 815-963, 839-1186, 953-1092, 981-1531, 997-1251, 1003-1336, 1006-1537, 1007-1300, 1016-1323, 1023-1694, 1024-1240, 1026-1438, 1036-1295, 1038-1537, 1040-1291, 1049-1587, 1069-1500, 1075-1295, 1081-1699, 1084-1324, 1090-1360, 1098-1344, 1103-1699, 1105-1641, 1106-1645, 1124-1584, 1124-1673, 1131-1771, 1133-1333, 1138-1311, 1139-1422, 1157-1740, 1165-1304, 1172-1403, 1174-1678, 1176-1830, 1179-1466, 1180-1851, 1182-1398, 1183-1442, |
| | 1184-1505, 1189-1483, 1190-1628, 1192-1759, 1197-1478, 1206-1702, 1210-1768, 1210-1891, 1215-1491, 1218-1600, 1220-1710, 1220-1817, 1226-1813, 1238-1701, 1239-1549, 1240-1872, 1247-1724, 1253-1506, 1256-1432, 1256-1588, 1256-1932, 1257-1861, 1268-1590, 1271-1593, 1272-1801, 1273-1544, 1275-1954, 1277-1900, 1288-1982, 1290-1489, 1295-1478, 1305-1733, 1309-1733, 1311-1812, 1312-1609, 1313-1982, 1315-1782, 1327-1454, 1329-1973, 1333-1951, 1335-1913, 1345-1949, 1348-1900, 1368-1952, 1374-1982, 1375-1982, 1377-1552, 1384-1666, 1389-1975, 1395-1982, 1397-1696, 1400-1672, 1400-1676, 1401-1785, 1401-1961, 1413-1748, 1413-1874, 1415-1847, 1420-1671, 1421-1723, 1430-1979, 1445-1662, 1445-1665, 1446-1965, 1466-1982, 1476-1982, 1482-1893, 1483-1972, 1483-1977, 1485-1971, 1486-1601, 1486-1737, 1494-1955, 1496-1972, 1508-1682, 1508-1726, 1508-1972, 1509-1977, 1513-1982, 1521-1982, 1523-1643, 1534-1972, 1534-1978, 1535-1976, 1538-1972, 1539-1977, 1540-1982, 1542-1776, 1543-1972, 1546-1978, 1548-1967, 1548-1972, 1550-1971, 1550-1972, 1561-1972, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 1578-1900, 1580-1972, 1584-1975, 1588-1972, 1590-1792, 1593-1972, 1598-1973, 1610-1976, 1613-1974, 1623-1897, 1625-1972, 1630-1971, 1636-1777, 1636-1925, 1638-1982, 1645-1753, 1646-1972, 1647-1891, 1648-1972, 1659-1972, 1662-1873, 1667-1972, 1669-1964, 1670-1978, 1671-1971, 1671-1974, 1672-1979, 1674-1972, 1678-1973, 1680-1975, 1682-1972, 1687-1973, 1702-1972, 1707-1971, 1709-1974, 1717-1973, 1717-1974, 1743-1978, 1768-1947, 1770-1973, 1779-1971, 1837-1974, 1838-1972, 1861-1982, 1899-1972 |
| 111/7505469CB1/2231 | 1-593, 4-307, 4-668, 4-753, 4-825, 4-2221, 32-606, 39-105, 91-392, 120-254, 126-410, 149-764, 156-756, 169-620, 178-684, 258-784, 271-717, 279-810, 289-789, 306-741, 402-665, 518-686, 533-686, 563-681, 665-779, 665-929, 665-936, 665-1058, 665-1095, 665-1116, 665-1297, 666-1155, 792-1131, 810-1039, 810-1418, 882-1241, 912-1463, 1038-1476, 1039-1633, 1073-1362, 1112-1560, 1125-1673, 1130-1401, 1135-1658, 1151-2012, 1167-1283, 1169-2012, 1174-2012, 1197-1673, 1204-1522, 1207-2012, 1208-2012, 1209-2012, 1212-2009, 1212-2012, 1226-2012, 1233-1673, 1242-1822, 1252-1758, 1253-2012, 1277-2012, 1296-1673, 1299-2012, 1335-2012, 1342-1745, 1344-2012, 1357-2012, 1368-2011, 1376-1673, 1382-2012, 1383-2012, 1390-2012, 1414-1712, 1423-2096, 1438-2099, 1450-1687, 1646-2221, 1668-2012, 1712-1952, 1712-1985, 1712-2192, 1712-2231, 1768-2231, 1780-2231, 1782-2231, 1814-2089, 1829-2231, 1840-2231, 1870-2004, 1880-2197, 1881-2231, 1900-2152, 1901-2145, 1909-2152, 1921-2231, 1928-2152, 1935-2153, 1950-2155, 1976-2147, 1982-2192, 2017-2231, 2034-2139 |
| 112/7505475CB1/5170 | 1-229, 1-255, 1-273, 1-420, 1-480, 1-507, 1-515, 1-563, 1-579, 1-594, 1-599, 1-630, 1-631, 1-636, 1-637, 1-5070, 4-314, 5-286, 14-275, 39-283, 43-500, 43-556, 51-320, 777-1350, 777-1377, 777-1394, 1130-1653, 1130-1836, 1230-1983, 1239-1525, 1239-1860, 1248-1852, 1274-1825, 1353-1923, 1359-1860, 1388-1987, 1432-1860, 1450-1979, 1450-1980, 1450-2029, 1450-2032, 1450-2066, 1450-2081, 1457-1860, 1459-1860, 1465-1860, 1475-1860, 1476-1964, 1527-1753, 1610-2048, 1615-2034, 1730-1908, 1838-2486, 1898-2343, 1898-2344, 2086-2243, 2091-2343, 2096-2286, 2096-2343, 2096-2344, 2102-2236, 2103-2236, 2104-2236, 2122-2411, 2187-2344, 2196-2769, 2196-2829, 2275-2673, 2275-2708, 2275-2767, 2275-2845, 2275-2851, 2275-2860, 2275-2879, 2275-2894, 2275-2912, 2275-2969, 2371-2778, 2389-2834, 2402-2943, 2405-2642, 2428-2767, 2450-2968, 2470-3070, 2476-2944, 2521-3218, 2612-3230, 2626-3124, 2626-3152, 2626-3184, 2626-3194, 2626-3211, 2626-3228, 2626-3247, 2626-3288, 2628-3115, 2629-3142, 2632-3206, 2632-3244, 2632-3356, 2632-3359, 2667-3361, 2679-3244, 2680-2948, |

Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 2680-3165, 2680-3227, 2680-3244, 2680-3361, 2682-3359, 2684-3236, 2695-3353, 2697-3361, 2701-3361, 2702-3348, 2703-3309, 2713-3361, 2717-3361, 2722-3361, 2727-3361, 2728-3361, 2730-3361, 2732-3361, 2733-3189, 2733-3317, 2733-3323, 2733-3361, 2738-3361, 2741-2985, 2741-3022, 2741-3132, 2741-3135, 2741-3199, 2741-3261, 2741-3273, 2741-3346, 2741-3361, 2742-3361, 2744-3041, 2744-3361, 2746-3361, 2752-3361, 2758-3361, 2764-3361, 2782-3283, 2782-3315, 2782-3357, 2782-3360, 2782-3361, 2784-3361, 2786-3361, 2789-3361, 2790-3361, 2806-3360, 2806-3361, 2829-3361, 2832-3360, 2833-3314, 2833-3333, 2833-3360, 2833-3361, 2835-3361, 2837-3361, 2843-3334, 2843-3361, 2844-3361, 2858-3361, 2859-3361, 2875-3356, 2883-3361, 2890-3360, 2895-3361, 2912-3361, 2920-3361, 2954-3361, 3072-3361, 3096-3361, 3118-3360, 3118-3361, 3119-3361, 3129-3351, 3129-3544, 3137-3360, 3137-3361, 3199-3474, 3199-3645, 3199-3672, 3199-3759, 3199-3781, 3199-3801, |
| | 3199-3859, 3199-3861, 3268-3361, 3368-3751, 3368-3843, 3368-3945, 3368-3951, 3391-4426, 3401-3590, 3438-3916, 3444-4086, 3449-3932, 3481-3865, 3511-4397, 3521-3768, 3538-4155, 3541-4156, 3550-4044, 3553-3847, 3553-4145, 3588-4259, 3623-4288, 3652-4353, 3674-4356, 3683-4169, 3684-4183, 3687-4013, 3688-4180, 3695-4273, 3696-4022, 3703-3952, 3713-4227, 3715-4288, 3727-4262, 3731-4406, 3734-4341, 3737-4285, 3739-4373, 3754-3991, 3758-4221, 3758-4321, 3758-4420, 3762-4313, 3763-4319, 3769-4202, 3772-4325, 3772-4402, 3796-4284, 3802-4471, 3817-4412, 3824-4341, 3825-4095, 3844-4407, 3846-4256, 3847-4365, 3848-4261, 3848-4299, 3848-4319, 3848-4353, 3848-4359, 3848-4427, 3849-4393, 3853-4203, 3857-4125, 3862-4109, 3863-4423, 3868-4434, 3869-4130, 3872-4369, 3878-4552, 3879-4399, 3887-4334, 3888-4621, 3894-4558, 3896-4483, 3901-4519, 3902-4165, 3915-4577, 3921-4564, 3925-4539, 3927-4137, 3930-4543, 3935-4177, 3940-4410, 3941-4249, |
| | 3953-4465, 3955-4416, 3955-4529, 3958-4146, 3960-4543, 3966-4651, 3967-4470, 3967-4598, 3970-4732, 3977-4652, 3978-4461, 3983-4652, 3988-4518, 3989-4545, 3991-4431, 3991-4520, 3994-4560, 3997-4648, 4004-4480, 4012-4272, 4012-4513, 4012-4709, 4019-4630, 4024-4691, 4027-4677, 4028-4531, 4032-4505, 4033-4504, 4034-4670, 4037-4297, 4037-4320, 4037-4518, 4037-4537, 4037-4656, 4039-4513, 4039-4621, 4040-4394, 4052-4725, 4053-4397, 4061-4568, 4065-5025, 4069-4663, 4074-4765, 4075-4418, 4075-4693, 4077-4583, 4086-4208, 4101-4607, 4103-4599, 4106-4539, 4139-4254, 4145-4757, 4148-4784, 4150-4706, 4154-4375, 4159-4365, 4170-4362, 4170-4651, 4172-4737, 4180-4709, 4183-4884, 4216-4835, 4219-4737, 4219-4934, 4221-4834, 4223-4678, 4228-4847, 4229-4983, 4239-4858, 4244-4843, 4252-4812, 4262-4653, 4263-4514, 4267-4408, 4280-4731, 4285-4930, 4286-4458, 4288-4894, 4290-4833, 4292-4840, 4294-4940, 4311-4835, 4317-4736, 4317-4946, 4326-4929, 4328-4429, 4331-4909, 4354-4924, 4358-4924, 4370-4950, 4370-4991, 4371-4950, 4377-4950, 4378-4952, 4388-5078, |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| | 4437-4959, 4443-5115, 4462-4950, 4490-4800, 4490-5070, 4509-4705, 4509-4722, 4511-4996, 4515-4779, 4515-4989, 4516-4794, 4525-4950, 4537-4950, 4553-4847, 4564-4950, 4566-4950, 4569-4945, 4577-4841, 4579-5170, 4599-5018, 4622-4947, 4625-4771, 4639-5041, 4642-4945, 4657-4959, 4657-5076, 4687-4979, 4701-5040, 4734-4952, 4755-5040, 4757-4872, 4757-4911, 4828-5066, 4922-5072 |
| 113/7505568CB1/1876 | 1-204, 1-246, 1-576, 1-603, 1-750, 1-765, 1-783, 1-816, 1-867, 1-1876, 2-634, 5-685, 7-206, 10-244, 30-281, 48-115, 102-865, 142-836, 305-547, 305-571, 305-573, 305-875, 326-907, 342-1225, 349-1036, 368-882, 380-941, 381-879, 383-588, 383-943, 385-519, 387-1006, 407-612, 429-948, 437-952, 451-1273, 455-1062, 463-994, 470-938, 475-734, 475-1022, 484-1086, 513-1497, 514-1143, 520-1121, 522-1093, 531-1101, 544-847, 548-1022, 573-1079, 585-1025, 585-1095, 585-1444, 607-943, 617-940, 619-1155, 620-1176, 626-1273, 631-1404, 632-1138, 652-1234, 669-1184, 670-1185, 673-1182, 680-1223, 680-1242, 680-1303, 700-801, 700-926, 712-1313, 713-868, 713-899, 727-1394, 740-1012, 740-1157, 740-1230, 740-1300, 740-1319, 740-1321, 740-1350, 740-1383, 740-1399, 741-1475, 742-878, 742-909, 742-1069, 742-1400, 744-1571, 752-978, 756-1030, 756-1041, 756-1230, 756-1295, 766-1304, 771-1213, 775-1404, 778-1256, 779-979, 779-988, 789-1149, 799-1308, 806-1064, 808-1149, 821-1425, 825-1383, 832-1408, 834-1086, 836-1424, 873-1480, 876-1173, 894-1204, 912-1437, 915-1740, 918-1480, 940-1454, 951- |
| | 1049-1412, 1054-1876, 1055-1876, 1079-1876, 1091-1876, 1123-1343, 1129-1876, 1143-1257, 1160-1876, 1185-1345, 1186-1729, 1191-1691, 1195-1876, 1239-1874, 1265-1753, 1270-1479, 1270-1497, 1270-1854, 1279-1820, 1325-1497, 1495-1875 |
| 114/7506953CB1/1602 | 1-1602, 114-308, 114-876, 308-1118, 308-1122, 311-1122, 324-805, 325-1122, 333-1122, 337-1122, 339-1121, 339-1122, 352-1122, 355-1122, 358-1122, 379-1118, 397-1122, 399-1118, 400-1121, 412-1122, 414-1122, 419-1118, 436-823, 441-823, 446-1118, 447-1118, 450-1122, 452-1118, 456-1122, 497-1062, 500-1122, 531-1122, 603-1118, 604-966, 604-1118, 627-1122, 881-1370, 923-1333, 923-1416, 925-1180, 1138-1488 |
| 115/7510176CB1/2173 | 1-398, 1-569, 1-644, 1-656, 1-724, 1-729, 1-736, 18-115, 18-124, 18-2173, 112-736, 410-952, 576-1116, 656-1460, 683-1618, 691-1607, 933-1870, 1053-1977, 1173-1814, 1373-2173, 1449-2173, 1453-2173, 1463-2173, 1465-2173, 1471-2173, 1823-1985 |

# Table 4

| Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length | Sequence Fragments |
|---|---|
| 116/7510541CB1/1826 | 1-237, 1-1826, 40-464, 49-371, 58-301, 59-713, 61-313, 66-206, 67-382, 71-341, 464-946, 464-1028, 467-946, 492-911, 498-1102, 499-1102, 508-1102, 517-782, 522-845, 524-822, 547-685, 580-877, 608-882, 743-1384, 822-1056, 822-1154, 822-1294, 822-1315, 826-1258, 842-1456, 870-1418, 892-1379, 922-1448, 939-1580, 953-1228, 958-1466, 1074-1280, 1086-1364, 1086-1393, 1091-1659, 1137-1750, 1138-1706, 1146-1488, 1175-1442, 1180-1458, 1180-1586, 1180-1678, 1180-1703, 1180-1826, 1182-1406, 1200-1775, 1243-1680, 1247-1784, 1294-1458, 1368-1650, 1374-1516, 1389-1789, 1459-1666, 1488-1773, 1584-1826 |
| 117/7510923CB1/2052 | 1-237, 1-2052, 40-466, 49-371, 50-813, 51-664, 51-785, 58-301, 60-637, 61-313, 67-382, 71-341, 71-666, 73-665, 203-778, 436-1008, 532-1019, 566-985, 596-919, 621-759, 654-951, 682-956, 915-1464, 987-1584, 1093-1730, 1103-1378, 1108-1616, 1224-1430, 1236-1514, 1236-1543, 1241-1809, 1287-1900, 1288-1856, 1296-1638, 1325-1592, 1330-1608, 1330-1736, 1330-1828, 1330-1853, 1330-1981, 1332-1556, 1350-1925, 1393-1830, 1397-1934, 1444-1608, 1518-1800, 1524-1666, 1539-1939, 1609-1816, 1638-1923, 1734-2052 |
| 118/7510984CB1/5056 | 1-5053, 291-687, 1037-1504, 1251-1382, 1969-2443, 1978-2162, 2304-2666, 2457-2724, 2579-3381, 3662-4331, 3708-4226, 4331-4585, 4331-4810, 4341-4667, 4349-4591, 4350-4691, 4350-4812, 4368-4513, 4368-4667, 4368-4685, 4368-4771, 4368-4798, 4368-4805, 4368-4817, 4368-4819, 4368-4825, 4368-4841, 4375-4617, 4379-5053, 4399-4612, 4469-5027, 4613-5056, 4638-5056, 4683-5036, 4695-5053, 4769-5056, 4894-5049 |

# Table 5

| Polynucleotide SEQ ID NO: | Incyte Project ID: | Representative Library |
|---|---|---|
| 60 | 7509332CB1 | MONOTXN05 |
| 61 | 7509102CB1 | PROSTUT10 |
| 62 | 7509132CB1 | COLNNOT01 |
| 64 | 7509178CB1 | MUSCNOT11 |
| 65 | 7509214CB1 | DENDTNT01 |
| 66 | 7509244CB1 | MUSCDIT06 |
| 67 | 7509256CB1 | ISLTNOT01 |
| 68 | 7509395CB1 | MUSCNOT11 |
| 69 | 7503287CB1 | BRAUNOR01 |
| 70 | 7503320CB1 | BSTMNON02 |
| 71 | 7503335CB1 | BRATDIC01 |
| 73 | 7504530CB1 | BRSTNOT04 |
| 74 | 7509303CB1 | MIXDTME02 |
| 75 | 7509910CB1 | BRACDIK08 |
| 76 | 7509982CB1 | BRSTNOT01 |
| 77 | 7510082CB1 | LIVRNON08 |
| 78 | 7510367CB1 | BRAINON01 |
| 79 | 7510413CB1 | MCLDTXN05 |
| 80 | 1721303CB1 | SPLNNOT11 |
| 81 | 7502007CB1 | DRGTNON04 |
| 82 | 7506439CB1 | ADENINB01 |
| 84 | 7509404CB1 | ISLTNOT01 |
| 85 | 7509439CB1 | LIVRTMR01 |
| 86 | 7510202CB1 | HEALDIR01 |
| 87 | 7510203CB1 | FIBRTXS07 |
| 88 | 7510208CB1 | BRAUNOR01 |
| 89 | 7510446CB1 | BRAITUT21 |
| 90 | 7505294CB1 | CORPNOT02 |
| 91 | 7505631CB1 | HUVELPB01 |
| 93 | 7510733CB1 | NEUTGMT01 |
| 94 | 7510734CB1 | NEUTGMT01 |
| 95 | 7503977CB1 | PROSTUT12 |
| 96 | 7505084CB1 | SINTBST01 |
| 97 | 7506950CB1 | BRAIHCT01 |
| 98 | 7506951CB1 | BRAIHCT01 |
| 99 | 7506954CB1 | BRAIHCT01 |
| 100 | 7506956CB1 | BRAIHCT01 |
| 101 | 7506959CB1 | BRAINOR03 |
| 102 | 7506960CB1 | BRAIHCT01 |
| 103 | 7510540CB1 | SINJNOT03 |
| 104 | 7510545CB1 | LUNGNOT22 |
| 105 | 7510654CB1 | BRAINOT12 |
| 106 | 7510660CB1 | FIBRTXS07 |
| 107 | 7510661CB1 | FIBRTXS07 |
| 108 | 7510680CB1 | LNODNON02 |
| 109 | 7505145CB1 | HEAONOT04 |
| 110 | 7505162CB1 | PROSNOT06 |
| 111 | 7505469CB1 | UTRSTUE01 |
| 112 | 7505475CB1 | DRGCNOT01 |

# Table 5

| Polynucleotide SEQ ID NO: | Incyte Project ID: | Representative Library |
|---|---|---|
| 113 | 7505568CB1 | LIVRNON08 |
| 114 | 7506953CB1 | BRAINOR03 |
| 115 | 7510176CB1 | GBLADIT01 |
| 116 | 7510541CB1 | MLP000060 |
| 117 | 7510923CB1 | LIVRTMR01 |
| 118 | 7510984CB1 | LIVRTUT13 |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| ADENINB01 | PBLUESCRIPT | Library was constructed using RNA isolated from the inflamed adenoid tissue of a 3-year-old child. (RNA came from Clontech.) |
| BRACDIK08 | PSPORT1 | This amplified and normalized library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day for 40 years). |
| BRAIHCT01 | pINCY | Library was constructed using RNA isolated from diseased occipital lobe tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema. |
| BRAINON01 | PSPORT1 | Library was constructed and normalized from 4.88 million independent clones from a brain tissue library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebralmeningeal lesion. Pathology for the associated tumor tissue indicated agrade 4 oligoastrocytoma in the right fronto-parietal part of the brain. The normalization and hybridization conditions were adapted from Soares etal., PNAS (1994) 91:9228, except that a significantly longer (48-hour) reannealing hybridization was used. |
| BRAINOR03 | PBK-CMV | This random primed library was constructed using pooled cDNA from two donors. cDNA was generated using mRNA isolated from brain tissue removed from a Caucasian male fetus (donor A) who was stillborn with a hypoplastic left heart at 23 weeks' gestation and from brain tissue removed from a Caucasian male fetus (donor B), who died at 23 weeks' gestation from premature birth. Serologies were negative for both donors and family history for donor B included diabetes in the mother. |
| BRAINOT12 | pINCY | Library was constructed using RNA isolated from brain tissue removed from the right frontal lobe of a 5-year-old Caucasian male during ahemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), which are consistent with chronic seizure disorder. Family history included acervical neoplasm. |
| BRAITUT21 | pINCY | Library was constructed using RNA isolated from brain tumor tissue removed from the midline frontal lobe of a 61-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated subfrontal meningothelial meningioma with no atypia. One ethmoid and mucosal tissue sample indicated meningioma. Family history included cerebrovascular disease, senile dementia, hyperlipidemia, benign hypertension, atherosclerotic coronary artery disease, congestive heart failure, and breast cancer. |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| BRATDIC01 | pINCY | This large size-fractionated library was constructed using RNA isolated from diseased brain tissue removed from the left temporal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology for the left temporal lobe, including the mesial temporal structures, indicated focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. The left frontal lobe showed a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remoteperinatal injury. The frontal lobe tissue also showed mild to moderategeneralized gliosis, predominantly subpial and subcortical, consistent withchronic seizure disorder. GFAP was positive for astrocytes. The patient presented with intractable epilepsy, focal epilepsy, hemiplegia, and anunspecified brain injury. Patient history included cerebral palsy, abnormality of gait, depressive disorder, and tobacco abuse in remission. Previous surgeries included tendon transfer. Patient medications included minocycline, hydrochloride, Tegretol, phenobarbital, vitamin C, Pepcid, and Pevaryl. Family history included brain cancer |
| | | in the father. |
| BRAUNOR01 | pINCY | This random primed library was constructed using RNA isolated fromstriatum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderateatherosclerosis involving the internal carotids, bilaterally; microscopicinfarcts of the frontal cortex and hippocampus, and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The |
| | | posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease. |
| BRSTNOT01 | PBLUESCRIPT | Library was constructed using RNA isolated from the breast tissue of a 56-year-old Caucasian female who died in a motor vehicle accident. |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| BRSTNOT04 | PSPORT1 | Library was constructed using RNA isolated from breast tissue removed froma 62-year-old East Indian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated an invasive grade 3 ductal carcinoma. Patient history included benign hypertension, hyperlipidemia, and hematuria. Family history included cerebrovascular and cardiovascular disease, hyperlipidemia, and liver cancer. |
| BSTMNON02 | PSPORT1 | This normalized brain stem library was constructed from 2.84 million independent clones from a brain stem library. Starting RNA was made from the brain stem tissue of a 72-year-old Caucasian male who died from myocardial infarction. Patient history included coronary artery disease, insulin-dependent diabetes mellitus, and arthritis. Normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228). |
| COLNNOT01 | PSPORT1 | Library was constructed using RNA isolated from colon tissue removed from a 75-year-old Caucasian male during a hemicolectomy. |
| CORPNOT02 | pINCY | Library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease. |
| DENDTNT01 | pINCY | Library was constructed using RNA isolated from treated dendritic cells from peripheral blood. |
| DRGCNOT01 | pINCY | Library was constructed using RNA isolated from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema and bronchopneumonia, bilateral pleural and pericardial effusions, and malignant lymphoma (natural killer cell type). Patient history included probable cytomegalovirus infection, hepatic congestion and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, and Bell's palsy. Surgeries included colonoscopy, large intestine biopsy, adenotonsillectomy, and nasopharyngeal endoscopy andbiopsy; treatment included radiation therapy. |
| DRGTNON04 | pINCY | The normalized dorsal root ganglion tissue library was constructed from 5.64 million independent clones from the a dorsal root ganglion library. Starting RNA was made from thoracic dorsal root ganglion tissue from a 32-year-old Caucasian male, who died from acute pulmonary edema, acute bronchopneumonia, pleural and pericardial effusion, and lymphoma. The patient presented with pyrexia, fatigue, and GI bleeding. Patient history included probable cytomegalovirus infection, liver congestion andsteatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, respiratory failure, pneumonia, natural killer cell lymphoma of the pharynx, Bell'spalsy, and tobacco and alcohol abuse. The library was normalized in one round using conditions adapted from Soares et al., PNAS(1994) 91:9228 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48-hours/round) reannealing hybridization was used. The library was then linearized and recircularized to select for insert containing clones as follows: plasmid DNA was prepped from |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| | | approximately 1 million clones from the normalized dorsal root ganglion tissue library following soft agar transformation. |
| FIBRTXS07 | pINCY | This subtracted library was constructed using 1.3 million clones from a dermal fibroblast library and was subjected to two rounds of subtraction hybridization with 2.8 million clones from an untreated dermal fibroblast tissue library. The starting library for subtraction was constructed using RNA isolated from treated dermal fibroblast tissue removed from the breast of a 31-year-old Caucasian female. The cells were treated with 9 CIS retinoic acid. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated dermal fibroblast tissue from the same donor. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR (1991) 19:1954 and Bonaldo, et al., Genome Research (1996) 6:791. |
| GBLADIT01 | pINCY | The library was constructed using RNA isolated from diseased gallbladder tissue removed from a 18-year-old Caucasian female during cholecystectomy and incidental appendectomy. Pathology indicated acute and chronic cholecystitis with cholelithiasis. The gallbladder contained multiple fragments of stony material. The appendix showed lymphoid hyperplasia. The patient presented with abdominal pain, nausea, and vomiting. Patienthistory included Chlamydia, extrinsicasthma, and cesarean delivery (x3). Family history included benign hypertension, acute myocardial infarction, and atherosclerotic coronary artery disease. |
| HEALDIR01 | PCDNA2.1 | This random primed library was constructed using RNA isolated from diseased left ventricle tissue removed from a 7-month old Caucasian male who died from cardiopulmonary arrest due to Pompe's disease. Patient history included Pompe's disease, left ventricular hypertrophy, pyrexia, right complete cleft lip, cleft palate, chronic serous otitis media, hypertrophic cardiomyopathy, congestive heart failure, and developmental delays. Family history included acute myocardial infarction, diabetes, cystic fibrosis and Down's syndrome. |
| HEAONOT04 | pINCY | Library was constructed using RNA isolated from aortic tissue removed from a 12-year-old Caucasian female, who died from a closed head injury. |
| HUVELPB01 | PBLUESCRIPT | Library was constructed using RNA isolated from HUV-EC-C (ATCC CRL 1730) cells that were stimulated with cytokine/LPS. RNA was isolated from two pools of HUV-EC-C cells that had been treated with either gamma IFN andTNF-alpha or IL-1 beta and LPS. In the first instance, HUV-EC-C cells were treated with 4 units/ml TNF and 2 units/ml IFNg for 96 hours. In the second instance, cells were treated with 1 units/ml IL-1 and 100 ng/ml LPS for 5 hours. |
| ISLTNOT01 | pINCY | Library was constructed using RNA isolated from a pooled collection of pancreatic islet cells. |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| LIVRNON08 | pINCY | This normalized library was constructed from 5.7 million independent clones from a pooled liver tissue library. Starting RNA was made from pooled liver tissue removed from a 4-year-old Hispanic male who died from anoxia and a 16 week female fetus who died after 16-weeks gestation from anencephaly. Serologies were positive for cytolomegalovirus in the 4-year-old. Patient history included asthma in the 4-year-old. Family history included taking daily prenatal vitamins and mitral valve prolapsein the mother of the fetus. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo etal., Genome Research 6 (1996):791, except that a significantly longer (48hours/round) reannealing hybridization was used. |
| LIVRTMR01 | PCDNA2.1 | This random primed library was constructed using RNA isolated from liver tissue removed from a 62-year-old Caucasian female during partial hepatectomy and exploratory laparotomy. Pathology for the matched tumor tissue indicated metastatic intermediate grade neuroendocrine carcinoma, consistent with islet cell tumor, forming nodules ranging in size, in the lateral and medial left liver lobe. The pancreas showed fibrosis, chronic inflammation and fat necrosis consistent with pseudocyst. The gallbladder showed mild chronic cholecystitis. Patient history included malignant neoplasm of the pancreas tail, pulmonary embolism, hyperlipidemia, thrombophlebitis, joint pain in multiple joints, type II diabetes, benign hypertension, cerebrovascular disease, and normal delivery. Previous surgeries included distal pancreatectomy, total splenectomy, and partial hepatectomy. Family history included pancreas cancer with secondary liver cancer, benign hypertension, and hyperlipidemia. |
| LIVRTUT13 | pINCY | Library was constructed using RNA isolated from liver tumor tissue removed from a 62-year-old Caucasian female during partial hepatectomy and exploratory laparotomy. Pathology indicated metastatic intermediate grade neuroendocrine carcinoma, consistent with islet cell tumor, forming nodules ranging in size, in the lateral and medial left liver lobe. The pancreas showed fibrosis, chronic inflammation and fat necrosis consistent with pseudocyst. The gall bladder showed mild chronic cholecystitis. Patienthistory included malignant neoplasm of the pancreas tail, pulmonary embolism, hyperlipidemia, thrombophlebitis, joint pain in multiple joints, type II diabetes, benign hypertension, and cerebrovascular disease. Family history included pancreas cancer, secondary liver cancer, benign hypertension, and hyperlipidemia. |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| LNODNON02 | pINCY | This normalized lymph node tissue library was constructed from .56 million independent clones from a lymph node tissue library. Starting RNA was made from lymph node tissue removed from a 16-month-old Caucasian male who died from head trauma. Serologies were negative. Patient history included bronchitis. Patient medications included Dopamine, Dobutamine, Vancomycin,Vasopressin, Proventil, and Atarax. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994)91:9228-9932 and Bonaldo et al., Genome Research 6 (1996):791, except thata significantly longer (48 hours/round) reannealing hybridization was used. |
| LUNGNOT22 | pINCY | Library was constructed using RNA isolated from lung tissue removed from a 58-year-old Caucasian female. The tissue sample used to construct this library was found to have tumor contaminant upon microscopic examination. Pathology for the associated tumor tissue indicated a caseating granuloma. Family history included congestive heart failure, breast cancer, secondary bone cancer, acute myocardial infarction and atherosclerotic coronary artery disease. |
| MCLDTXN05 | pINCY | This normalized dendritic cell library was constructed from 1 millionin dependent clones from a pool of two derived dendritic cell libraries.Starting libraries were constructed using RNA isolated from untreated andtreated derived dendritic cells from umbilical cord blood CD34+ precursor cells removed from a male. The cells were derived with granulocyte/macrophage colony stimulating factor (GM-CSF), tumor necrosis factor alpha (TNF alpha), and stem cell factor (SCF). The GM-CSF was added at time 0 at 100 ng/ml, the TNF alpha was added at time 0 at 2.5 ng/ml, and the SCF was added at time 0 at 25 ng/ml. Incubation time was 13 days. The treated cells were then exposed to phorbol myristate acetate (PMA), and Ionomycin. The PMA and Ionomycin were added at 13 days for five hours. The library was normalized in two rounds using conditions adapted from Soareset al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research 6(1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used. |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| MIXDTME02 | PBK-CMV | This 5' biased random primed library was constructed using pooled cDNA from five donors. cDNA was generated using mRNA isolated from heart tissue removed from a Caucasian male fetus who died after 20 weeks gestation from Patau's syndrome (donor A); adrenal gland removed from a 43-year-old Caucasian male (donor B) during nephroureterectomy, regional lymph nodeexcision and unilateral adrenalectomy; kidney cortex removed from a 65-year-old male (donor C) during nephroureterectomy; lung tissue removed from a 14-month-old Caucasian female who died from drowning (donor D); and kidney tissue removed from an 8-year-old Caucasian female who died from a motor vehicle accident (donor E). For donor B, pathology for the associated tumor indicated grade 2 (of 4) renal cell carcinoma in the left kidney with invasion into the renal pelvis. Patient presented with hematuria andanemia. Patient history included benign hypertension and obesity. Previous surgeries included adenotonsillectomy and indirect inguinal hernia repair. The patient was |
| | | not taking any medications. Family history included benign hypertension and atherosclerotic coronary artery disease in the father. For donor C pathology for the associated tumor shows grade 3 (of 4) renal cell carcinoma, clear cell type, within the mid-portion of the kidney. For donor D, serologies were negative. For donor E, medications included respiradol. MLP000060 PCR2-TOPOTA Library was constructed using pooled cDNA from different donors. cDNA was generated using mRNA isolated from the following: aorta, cerebellum, lymphnodes, muscle, tonsil (lymphoid hyperplasia), bladder tumor (invasive grade 3 transitional cell carcinoma), breast (proliferative fibrocystic changes without atypia characterized by epithelial ductal hyperplasia, testicle tumor (embryonal carcinoma), spleen, ovary, parathyroid, ileum, breast skin, sigmoid colon, penis tumor (fungating invasive grade 4 squamous cellcarcinoma), fetal lung, breast, fetal small intestine, fetal liver, fetal pancreas, fetal lung, fetal skin, fetal penis, fetal bone, fetal ribs, frontal brain tumor (grade 4 |
| | | gemistocytic astrocytoma), ovary (stromal hyperthecosis), bladder, bladder tumor (invasive grade 3 transitional cell carcinoma), stomach, lymph node tumor (metastatic basaloid squamous cell carcinoma), tonsil (reactive lymphoid hyperplasia), periosteum from the tibia, fetal brain, fetal spleen, uterus tumor, endometrial (grade 3 adenosquamous carcinoma), seminal vesicle, liver, aorta, adrenal gland, lymph node (metastatic grade 3 squamous cell carcinoma), glossal muscle,esophagus, esophagus tumor (invasive grade 3 adeno carcinoma), ileum, pancreas, soft tissue tumor from the skull (grade 3 ependymoma), transverse colon, (benign familial polyposis), rectum tumor (grade 3 colonic adenocarcinoma), rib tumor, (metastatic grade 3 osteosarcoma), lung, heart, placenta, thymus, stomach, spleen (splenomegaly with congestion), uterus, cervix (mild chronic cervicitis with focal squamous metaplasia), spleen tumor (malignant lymphoma, diffuse large cell type, B-cell phenotype with abundant reactive T-cells and marked granulomatous response), umbilicalcord blood mononuclear cells, |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| | | upper lobe lung tumor, (grade 3 squamous cell carcinoma), endometrium (secretory phase), liver, liver tumor (metastatic grade 2 neuroendocrine carcinoma), colon, umbilical cord blood, Th1 cells, nonactivated, umbilical cord blood, Th2 cells, nonactivated, coronary artery endothelial cells (untreated), coronary artery smooth muscle cells, (untreated), coronary artery smooth muscle cells (treated with TNF & IL-110ng/ml each for 20 hours), bladder (mild chronic cystitis), epiglottis, breast skin, small intestine, fetal prostate stroma fibroblasts, prostate epithelial cells (PrEC cells), fetal adrenal glands, fetal liver, kidney transformed embryonal cell line (293-EBNA) (untreated), kidney transformed embryonal cell line (293-EBNA) (treated with 5Aza-2deoxy cytidine for 72 hours), mammary epithelial cells, (HMEC cells), peripheral blood monocytes (treated with IL-10 at time 0, 10ng/ml, LPS was added at 1 hour at 5ng/ml.Incubation 24 hours), peripheral blood monocytes (treated with anti-IL-10at time 0, 10ng/ml, LPS was added at 1 hour at 5ng/ml. Incubation 24 hours), spinal cord, base of |
| | | medulla (Huntington's chorea), thigh and arm muscle (ALS), breast skin fibroblast (untreated), breast skin fibroblast (treated with 9CIS Retinoic Acid 1$\mu$M for 20 hours), breast skin fibroblast (treated with TNF-alpha & IL-1 beta, 10ng/ml each for 20 hours), fetal liver mast cells, hematopoietic (Mast cells prepared from human fetal liver hematopoietic progenitor cells (CD34+ stem cells) cultured in the presence of hIL-6 and hSCF for 18 days), epithelial layer of colon, bronchial epithelial cells (treated for 20 hours with 20% smoke conditioned media), lymph node, pooled peripheral blood mononuclear cells (untreated), pooled brain segments: striatum, globus pallidus and posterior putamen (Alzheimer's Disease), pituitary gland, umbilical cord blood, CD34+ deriveddendritic cells (treated with SCF, GM-CSF & TNF alpha, 13 days), umbilical cord blood, CD34+ derived dendritic cells (treated with SCF, GM-CSF & TNFalpha, 13 days followed by PMA/Ionomycin for 5 hours), small intestine rectum, bone marrow neuroblastoma cell line (SH-SY5Y cells, treated with 6-Hydroxydopamine |
| | | 100 uM for 8 hours), bone marrow, neuroblastoma cell line (SH-SY5Y cells, untreated), brain segments from one donor: amygdala, entorhinal cortex, globus pallidus, substantia innominata, striatum, dorsalcaudate nucleus, dorsal putamen, ventral nucleus accumbens, archaecortex (hippocampus anterior and posterior), thalamus, nucleus raphe magnus, periaqueductal gray, midbrain, substantia nigra, and dentate nucleus, pineal gland (Alzheimer's Disease), preadipocytes (untreated), preadipocytes (treated with a peroxisome proliferator-activated receptor gamma agonist, 1microM, 4 hours), pooled prostate (adenofibromatoushyperplasia), pooled kidney, pooled adipocytes (untreated), pooledadipocytes (treated with human insulin), pooled mesentaric and abdomenal fat, pooled adrenal glands, pooled thyroid (normal and adenomatous hyperplasia), pooled spleen (normal and with changes consistent with idiopathic thrombocytopenic purpura), pooled right and left breast, pooledlung, pooled nasal polyps, pooled fat, pooled synovium (normal andrhumatoid arthritis), pooled brain |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| | | (meningioma, gemistocytic astrocytoma and Alzheimer's disease), pooled fetal colon, pooled colon: ascending, descending (chronic ulcerative colitis), and rectal tumor (adenocarcinoma), pooled esophagus, normal and tumor (invasive grade 3 adenocarcinoma), pooled breast skin fibroblast (one treated w/ 9CIS Retinoic Acid and theother with TNF-alpha & IL-1 beta), pooled gallbladder (acute necrotizingcholecystitis with cholelithiasis (clinically hydrops), acute hemorrhagic cholecystitis with cholelithiasis, chronic cholecystitis and cholelithiasis), pooled fetal heart, (Patau's and fetal demise), pooled neurogenic tumor cell line, SK-N-MC, (neuroepitelioma, metastasis tosupra-orbital area, untreated) and neuron, NT-2 cell line, (treated with mouse leptin at 1 $\mu$g/ml and 9cis retinoic acid at 3.3 $\mu$M for 6 days), pooled ovary (normal and polycystic ovarian disease), pooled prostate, (adenofibromatous hyperplasia), pooled seminal vesicle, pooled smal lintestine, pooled fetal small intestine, pooled stomach and fetal stomach, prostate epithelial cells, pooled |
| | | testis (normal and embryonal carcinoma), pooled uterus, pooled uterus tumor (grade 3 adenosquamous carcinoma andleiomyoma), pooled uterus, endometrium, and myometrium, (normal andadenomatous hyperplasia with squamous metaplasia and focal atypia), pooled brain: (temporal lobe meningioma, cerebellum and hippocampus (Alzheimer's Disease), pooled skin, fetal lung, adrenal tumor (adrenal cortical carcinoma), prostate tumor (adenocarcinoma), fetal heart, fetal smal lintestine, ovary tumor (mucinous cystadenoma), ovary, ovary tumor (transitional cell carcinoma), disease prostate (adeno fibromatous hyperplasia), fetal colon, uterus tumor (leiomyoma), temporal brain, submandibular gland, colon tumor (adenocarcinoma), ascending and transverse colon, ovary tumor (endometrioid carcinoma), lung tumor (squamous cell carcinoma), fetal brain, fetal lung, ureter tumor (transitional cell carcinoma), untreated HNT cells, para-aortic soft tissue, testis, seminal vesicle, diseased ovary (endometriosis), temporal lobe, myometrium, diseased gallbladder |
| | | (cholecystitis, cholelithiasis), placenta, breast tumor (ductal adenocarcinoma), breast, lung tumor (lipo sarcoma), endometrium, abdominal fat, cervical spine dorsal root ganglion, thoracic spine dorsal root ganglion, diseased thyroid (adenomatous hyperplasia), liver, kidney, fetal liver, NT-2 cells (treated with mouse leptin and 9 cisRA), K562 cells (treated with 9 cis RA), cerebellum, corpus callosum, hypothalamus, fetal brain astrocytes (treated with TNFa and IL-1b), inferior parietal cortex, posterior hippocampus, pons, thalamus, C3A cells (untreated), C3A cells (treated with 3-methylcholanthrene), testis, colon epithelial layer, pooled prostate, pooled liver, substantia nigra, thigh muscle, rib bone, fallopian tube tumor (endometrioid and serousadenocarcinoma), diseased lung (idiopathic pulmonary disease), cingulateanterior allocortex and neocortex, cingulate posterior allocortex, auditory neocortex, frontal neocortex, orbital inferior neocortex, parietal superior neocortex, visual primary neocortex, dentate nucleus, posterior cingulate, cerebellum, vermis, |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| | | inferior temporal cortex, medulla, posterior parietal cortex, colon polyp, pooled breast, anterior and posterior hippocampus, mesenteric and abdominal fat, pooled esophagus, pooled fetal kidney, pooled fetal liver, ileum, small intestine, pooled gallbladder, frontal and superior temporal cortex, pooled ovary, pooled endometrium, pooled prostate, pooled kidney, fetal femur, sacrum tumor (giant cell tumor), pooled kidney and kidney tumor (renal cell carcinoma clear-cell type), pooled liver and liver tumor (neuroendocrine carcinoma), pooled fetal liver, pooled lung, fetal pancreas, pancreas, parotid gland, parotid tumor (sebaceous lymphadenoma), retroperitoneal and suprglottic soft tissue, spleen, fetal spleen, spleen tumor (malignant lymphoma), diseased spleen (idiopathic thrombocytopenic purpura), parathyroid, thyroid, thymus, tonsil ureter tumor (transitional cell carcinoma), pooled adrenal gland and adrenal tumor (pheochromocytoma), pooled lymph node tumor (Hodgkin's disease and metastatic adenocarcinoma), pooled neck and calf muscles, and pooled bladder. |
| MONOTXN05 | pINCY | This normalized treated monocyte cell tissue library was constructed from 1.03 million independent clones from a monocyte tissue library. Starting RNA was made from RNA isolated from treated monocytes from peripheral blood removed from a 42-year-old female. The cells were treated with interleukin-10 (IL-10) and lipopolysaccharide (LPS). The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used. |
| MUSCDIT06 | pINCY | Library was constructed using RNA isolated from skeletal muscle tissue removed from an 11-month-old Caucasian female who died from cardiopulmonary arrest. Patient history included Krabbe's disease. |
| MUSCNOT11 | pINCY | The library was constructed using RNA isolated from diseased arm muscle tissue removed from a 74-year-old Caucasian female who died from respiratory arrest due to amyotrophic lateral sclerosis (ALS). Patient historyincluded amyotrophic lateral sclerosis, hypertension, arthritis, and alcohol use. |
| NEUTGMT01 | PSPORT1 | Library was constructed using RNA isolated from peripheral blood granulocytes collected by density gradient centrifugation through Ficoll-Hypaque. The cells were isolated from buffy coat units obtained from20 unrelated male and female donors. Cells were cultured in 10 nM GM-CSF for 1 hour before washing and harvesting for total RNA preparation. |
| PROSNOT06 | PSPORT1 | Library was constructed using RNA isolated from the diseased prostate tissue of a 57-year-old Caucasian male during radical prostatectomy, removal of both testes and excision of regional lymph nodes. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma (Gleason grade 3+3). Patient history included a benign neoplasm of the large bowel and type I diabetes. Family history included a malignant neoplasm of the prostate and type I diabetes. |

# Table 6

| Library | Vector | Library Description |
|---|---|---|
| PROSTUT10 | pINCY | Library was constructed using RNA isolated from prostatic tumor tissuere moved from a 66-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated an adenocarcinoma (Gleason grade 2+3). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer and secondary bone cancer. |
| PROSTUT12 | pINCY | Library was constructed using RNA isolated from prostate tumor tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated an adenocarcinoma (Gleason grade 2+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). |
| SINJNOT03 | pINCY | Library was constructed using RNA isolated from duodenum tissue removed from the small intestine of a 16-year-old Caucasian male who died from head trauma. Patient history included a kidney infection. |
| SINTBST01 | pINCY | Library was constructed using RNA isolated from ileum tissue obtained froman 18-year-old Caucasian female during bowel anastomosis. Pathology indicated Crohn's disease of the ileum, involving 15 cm of the small bowel. Family history included cerebrovascular disease and atherosclerotic coronary artery disease. |
| SPLNNOT11 | pINCY | Library was constructed using RNA isolated from diseased spleen tissue removed from a 14-year-old Asian male during a total splenectomy. Pathology indicated changes consistent with idiopathic thrombocytopenic purpura. The patient presented with bruising. Patient medications included Vincristine. |
| UTRSTUE01 | PCDNA2.1 | This 5' biased random primed library was constructed using RNA isolated from uterus tumor tissue removed a 37-year-old Black female during myomectomy, dilation and curettage, right fimbrial region biopsy, and incidental appendectomy. Pathology indicated multiple (12) uterine leiomyomata. A fimbrial cyst was identified. The patient presented with deficiency anemia, an umbilical hernia, and premenopausal menorrhagia. Patient history included premenopausal menorrhagia and sarcoidosis of the lung. Previous surgeries included hysteroscopy, dilation and curettage, and an endoscopic lung biopsy. Patient medications included Chromagen and Claritin. Family history included acute myocardial infarction and atherosclerotic coronary artery disease in the father. |

# Table 7

| Program | Description | Reference | Parameter Threshold |
|---|---|---|---|
| ABI FACTURA | A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences. | Applied Biosystems, Foster City, CA. | |
| ABI/PARACEL FDF | A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences. | Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA. | Mismatch <50% |
| ABI AutoAssembler | A program that assembles nucleic acid sequences. | Applied Biosystems, Foster City, CA. | |
| BLAST | A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx. | Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402. | ESTs: Probability value = 1.0E-8 or less; Full Length sequences: Probability value = 1.0E-10 or less |
| FASTA | A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch. | Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489. | ESTs: fasta E value = 1.06E-6; Assembled ESTs: fasta Identity = 95% or greater and Match length = 200 bases or greater; fastx E value = 1.0E-8 or less; Full Length sequences: fastx score = 100 or greater |
| BLIMPS | A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions. | Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424. | Probability value = 1.0E-3 or less |

# Table 7

| Program | Description | Reference | Parameter Threshold |
|---|---|---|---|
| HMMER | An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM, INCY, SMART and TIGRFAM. | Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350. | PFAM, INCY, SMART or TIGRFAM hits: Probability value = 1.0E-3 or less; Signal peptide hits: Score = 0 or greater |
| ProfileScan | An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite. | Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221. | Normalized quality score ≥ GCG specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1. |
| Phred | A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability. | Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194. | |
| Phrap | A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences. | Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA. | Score = 120 or greater; Match length = 56 or greater |
| Consed | A graphical tool for viewing and editing Phrap assemblies. | Gordon, D. et al. (1998) Genome Res. 8:195-202. | |
| SPScan | A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides. | Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439. | Score = 3.5 or greater |
| TMAP | A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation. | Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371. | |

# Table 7

| Program | Description | Reference | Parameter Threshold |
| --- | --- | --- | --- |
| TMHMMER | A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation. | Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. On Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence (AAAI) Press, Menlo Park, CA, and MIT Press, Cambridge, MA, pp. 175-182. | |
| Motifs | A program that searches amino acid sequences for patterns that matched those defined in Prosite. | Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI. | |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 60 | 7509332 | 1667982H1 | SNP00071142 | 118 | 319 | C | C | A | R28 | 1.00 | n/a | n/a | n/a |
| 62 | 7509132 | 7252817H2 | SNP00120204 | 200 | 144 | T | T | C | noncoding | n/a | n/a | n/a | n/a |
| 62 | 7509132 | 7252817J2 | SNP00131224 | 163 | 1787 | T | T | C | noncoding | n/a | n/a | n/a | n/a |
| 63 | 7509136 | 3000824H1 | SNP00005163 | 185 | 1524 | T | T | C | noncoding | 0.82 | 0.81 | 0.76 | 0.83 |
| 63 | 7509136 | 3002289H1 | SNP00005163 | 186 | 1523 | T | T | C | noncoding | 0.82 | 0.81 | 0.76 | 0.83 |
| 63 | 7509136 | 3506005H1 | SNP00005163 | 67 | 1525 | C | T | C | noncoding | 0.82 | 0.81 | 0.76 | 0.83 |
| 65 | 7509214 | 1305469H1 | SNP00057352 | 58 | 447 | G | A | G | A119 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 1305469H1 | SNP00136872 | 12 | 401 | G | G | A | R104 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 1305469H1 | SNP00136873 | 92 | 481 | C | C | T | P131 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 1637908H1 | SNP00136872 | 149 | 402 | G | G | A | R104 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 2129306H1 | SNP00047651 | 38 | 246 | G | G | T | M52 | n/d | n/a | n/a | n/a |
| 65 | 7509214 | 2371557H1 | SNP00136874 | 111 | 607 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 2371557H1 | SNP00004818 | 117 | 613 | G | T | G | noncoding | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 2517189H2 | SNP00136872 | 273 | 413 | G | G | A | R108 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 2734103H1 | SNP00057352 | 136 | 449 | G | A | G | W120 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 2734103H1 | SNP00136872 | 90 | 403 | A | G | A | N105 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 2734103H1 | SNP00136873 | 170 | 483 | C | C | T | P131 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 3603012H1 | SNP00136872 | 274 | 421 | G | G | A | G111 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 3837531H1 | SNP00136874 | 183 | 611 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 3837531H1 | SNP00057352 | 28 | 453 | G | A | G | G121 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 3837531H1 | SNP00136873 | 62 | 487 | C | C | T | R133 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 3946173H1 | SNP00136873 | 138 | 482 | C | C | T | P131 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 4008514H1 | SNP00047651 | 120 | 244 | T | G | T | L52 | n/d | n/a | n/a | n/a |
| 65 | 7509214 | 4108144H1 | SNP00047651 | 68 | 242 | G | G | T | G51 | n/d | n/a | n/a | n/a |
| 65 | 7509214 | 4386522H1 | SNP00057352 | 171 | 448 | G | A | G | G120 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 5605610H1 | SNP00004818 | 230 | 616 | G | T | G | noncoding | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 5966775H1 | SNP00136874 | 376 | 608 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 5966775H1 | SNP00004818 | 380 | 612 | G | T | G | noncoding | n/a | n/a | n/a | n/a |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 65 | 7509214 | 8611118H1 | SNP00004818 | 178 | 609 | G | T | G | noncoding | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 8611118H1 | SNP00057352 | 343 | 444 | G | A | G | L118 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 8611118H1 | SNP00136872 | 389 | 398 | G | G | A | R103 | n/a | n/a | n/a | n/a |
| 65 | 7509214 | 8611118H1 | SNP00136873 | 309 | 478 | C | C | T | R130 | n/a | n/a | n/a | n/a |
| 66 | 7509244 | 3000824H1 | SNP00005163 | 185 | 1491 | T | T | C | noncoding | 0.82 | 0.81 | 0.76 | 0.83 |
| 66 | 7509244 | 3002289H1 | SNP00005163 | 186 | 1490 | T | T | C | noncoding | 0.82 | 0.81 | 0.76 | 0.83 |
| 66 | 7509244 | 3506005H1 | SNP00005163 | 67 | 1492 | C | T | C | noncoding | 0.82 | 0.81 | 0.76 | 0.83 |
| 70 | 7503320 | 5681822H1 | SNP00014931 | 23 | 552 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 71 | 7503335 | 1454748H1 | SNP00072734 | 211 | 1691 | G | A | G | P536 | 0.43 | 0.45 | 0.14 | 0.22 |
| 71 | 7503335 | 1485166H1 | SNP00069424 | 237 | 310 | T | T | C | V76 | 0.93 | n/d | 0.96 | 0.96 |
| 71 | 7503335 | 2963710H1 | SNP00069425 | 138 | 1432 | C | A | C | A450 | 0.90 | 0.96 | 0.85 | 0.89 |
| 71 | 7503335 | 2963710H1 | SNP00072733 | 73 | 1367 | C | C | T | P428 | n/a | n/a | n/a | n/a |
| 71 | 7503335 | 4057086H1 | SNP00069424 | 39 | 308 | T | T | C | I75 | 0.93 | n/d | 0.96 | 0.96 |
| 71 | 7503335 | 4057086H1 | SNP00120820 | 176 | 445 | T | T | G | V121 | n/a | n/a | n/a | n/a |
| 71 | 7503335 | 4093467H1 | SNP00072733 | 217 | 1366 | C | C | T | P428 | n/a | n/a | n/a | n/a |
| 71 | 7503335 | 4295902H1 | SNP00072733 | 243 | 1365 | C | C | T | P428 | n/a | n/a | n/a | n/a |
| 71 | 7503335 | 6770662H1 | SNP00120820 | 405 | 447 | T | T | G | Y122 | n/a | n/a | n/a | n/a |
| 71 | 7503335 | 6867692H1 | SNP00120821 | 515 | 546 | C | C | T | H155 | 0.48 | 0.59 | 0.38 | 0.32 |
| 71 | 7503335 | 8084392H1 | SNP00126828 | 93 | 987 | G | A | G | A302 | n/a | n/a | n/a | n/a |
| 73 | 7504530 | 2639741H1 | SNP00119886 | 47 | 90 | A | A | C | noncoding | 0.37 | n/a | n/a | n/a |
| 73 | 7504530 | 3601392H1 | SNP00133389 | 204 | 824 | G | A | G | G207 | n/a | n/a | n/a | n/a |
| 73 | 7504530 | 4080023H1 | SNP00119886 | 18 | 87 | C | A | C | noncoding | 0.37 | n/a | n/a | n/a |
| 74 | 7509303 | 1546672H1 | SNP00075644 | 35 | 1320 | T | T | C | noncoding | n/a | n/a | n/a | n/a |
| 74 | 7509303 | 4112747H1 | SNP00005993 | 51 | 1818 | G | G | A | noncoding | 0.63 | 0.90 | 0.47 | 0.67 |
| 74 | 7509303 | 4897687H1 | SNP00005993 | 103 | 1827 | A | G | A | noncoding | 0.63 | 0.90 | 0.47 | 0.67 |
| 74 | 7509303 | 5732638H1 | SNP00005993 | 35 | 1828 | A | G | A | noncoding | 0.63 | 0.90 | 0.47 | 0.67 |
| 74 | 7509303 | 5955401H1 | SNP00005993 | 35 | 1825 | G | G | A | noncoding | 0.63 | 0.90 | 0.47 | 0.67 |
| 74 | 7509303 | 6412945H1 | SNP00005993 | 131 | 1810 | G | G | A | noncoding | 0.63 | 0.90 | 0.47 | 0.67 |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 74 | 7509303 | 6937415H1 | SNP00005993 | 388 | 1830 | G | G | A | noncoding | 0.63 | 0.90 | 0.47 | 0.67 |
| 75 | 7509910 | 1454748H1 | SNP00072734 | 211 | 1913 | G | A | G | noncoding | 0.43 | 0.45 | 0.14 | 0.22 |
| 75 | 7509910 | 1485166H1 | SNP00069424 | 237 | 310 | T | T | C | V76 | 0.93 | n/d | 0.96 | 0.96 |
| 75 | 7509910 | 2963710H1 | SNP00069425 | 138 | 1654 | C | A | C | noncoding | 0.90 | 0.96 | 0.85 | 0.89 |
| 75 | 7509910 | 2963710H1 | SNP00072733 | 73 | 1589 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 75 | 7509910 | 4057086H1 | SNP00069424 | 39 | 308 | T | T | C | I75 | 0.93 | n/d | 0.96 | 0.96 |
| 75 | 7509910 | 4057086H1 | SNP00120820 | 176 | 445 | T | T | G | V121 | n/a | n/a | n/a | n/a |
| 75 | 7509910 | 4093467H1 | SNP00072733 | 217 | 1588 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 75 | 7509910 | 4295902H1 | SNP00072733 | 243 | 1587 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 75 | 7509910 | 6770662H1 | SNP00120820 | 405 | 447 | T | T | G | Y122 | n/a | n/a | n/a | n/a |
| 75 | 7509910 | 6867692H1 | SNP00120821 | 515 | 546 | C | C | T | H155 | 0.48 | 0.59 | 0.38 | 0.32 |
| 75 | 7509910 | 7130883H1 | SNP00126828 | 9 | 1209 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 76 | 7509982 | 1307948H1 | SNP00023319 | 77 | 4204 | A | G | A | N1328 | 0.57 | 0.28 | 0.36 | 0.65 |
| 76 | 7509982 | 1307948H1 | SNP00023320 | 201 | 4328 | C | C | T | N1369 | n/d | n/d | n/d | n/d |
| 76 | 7509982 | 1307948H1 | SNP00072137 | 91 | 4218 | A | A | G | S1333 | n/a | n/a | n/a | n/a |
| 76 | 7509982 | 1965082H1 | SNP00023319 | 146 | 4203 | G | G | A | G1328 | 0.57 | 0.28 | 0.36 | 0.65 |
| 76 | 7509982 | 1965082H1 | SNP00072137 | 160 | 4217 | A | A | G | K1332 | n/a | n/a | n/a | n/a |
| 76 | 7509982 | 2764315H1 | SNP00026116 | 169 | 4072 | T | T | C | I1284 | n/d | n/a | n/a | n/a |
| 76 | 7509982 | 3781324H1 | SNP00026116 | 174 | 4069 | C | T | C | T1283 | n/d | n/a | n/a | n/a |
| 76 | 7509982 | 3781324H1 | SNP00139248 | 272 | 4167 | T | C | T | F1316 | n/a | n/a | n/a | n/a |
| 76 | 7509982 | 3948943H1 | SNP00023320 | 89 | 4327 | C | C | T | T1369 | n/d | n/d | n/d | n/d |
| 76 | 7509982 | 415352H1 | SNP00058466 | 177 | 5510 | C | C | G | noncoding | 0.62 | n/a | n/a | n/a |
| 76 | 7509982 | 4776786H1 | SNP00023319 | 67 | 4202 | G | G | A | P1327 | 0.57 | 0.28 | 0.36 | 0.65 |
| 76 | 7509982 | 4776786H1 | SNP00072137 | 81 | 4216 | A | A | G | K1332 | n/a | n/a | n/a | n/a |
| 76 | 7509982 | 5704531H1 | SNP00023320 | 113 | 4325 | T | C | T | D1368 | n/d | n/d | n/d | n/d |
| 76 | 7509982 | 5704531H1 | SNP00072137 | 3 | 4215 | A | A | G | K1332 | n/a | n/a | n/a | n/a |
| 79 | 7510413 | 1667982H1 | SNP00071142 | 118 | 319 | C | C | A | R28 | 1.00 | n/a | n/a | n/a |
| 79 | 7510413 | 2240820H2 | SNP00066386 | 8 | 358 | A | A | G | K41 | n/a | n/a | n/a | n/a |

Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 79 | 7510413 | 6844294H1 | SNP00119027 | 246 | 280 | A | A | G | M15 | n/d | n/d | n/d | n/d |
| 80 | 1721303 | 2905327H1 | SNP00095824 | 93 | 94 | T | T | C | P27 | n/a | n/a | n/a | n/a |
| 80 | 1721303 | 3881018H1 | SNP00095824 | 62 | 96 | T | T | C | I28 | n/a | n/a | n/a | n/a |
| 80 | 1721303 | 5765782H1 | SNP00095824 | 98 | 98 | C | T | C | R29 | n/a | n/a | n/a | n/a |
| 80 | 1721303 | 940621H1 | SNP00016445 | 89 | 233 | T | T | C | noncoding | n/a | n/a | n/a | n/a |
| 81 | 7502007 | 3601392H1 | SNP00133389 | 204 | 680 | G | A | G | G206 | n/a | n/a | n/a | n/a |
| 82 | 7506439 | 7753512J1 | SNP00039567 | 169 | 1803 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 82 | 7506439 | 7753512J1 | SNP00112674 | 25 | 1947 | T | C | T | noncoding | 0.97 | n/a | n/a | n/a |
| 85 | 7509439 | 015688H1 | SNP00139297 | 11 | 279 | C | C | T | A66 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 077004H1 | SNP00139297 | 199 | 281 | C | C | T | R67 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 1273186H1 | SNP00139297 | 216 | 280 | C | C | T | A66 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 1551917H1 | SNP00139297 | 185 | 276 | C | C | T | P65 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 167737H1 | SNP00139297 | 206 | 277 | C | C | T | P65 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 1844902H1 | SNP00139297 | 62 | 278 | C | C | T | P66 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 2866273H1 | SNP00139297 | 159 | 229 | C | C | T | N49 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 2960845H1 | SNP00139297 | 49 | 275 | C | C | T | P65 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 3026771H1 | SNP00139297 | 210 | 274 | C | C | T | N64 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 3080978H1 | SNP00139297 | 198 | 264 | C | C | T | T61 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 3115228H1 | SNP00139297 | 265 | 272 | C | C | T | Q64 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 3240634H1 | SNP00139297 | 227 | 273 | C | C | T | T64 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 4063887H1 | SNP00139297 | 36 | 271 | C | C | T | L63 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 6096843H1 | SNP00139297 | 218 | 267 | C | C | T | A62 | n/a | n/a | n/a | n/a |
| 85 | 7509439 | 6749571H1 | SNP00124031 | 40 | 45 | A | A | G | noncoding | n/d | n/a | n/a | n/a |
| 86 | 7510202 | 1005109H1 | SNP00004180 | 33 | 4131 | G | C | G | P1377 | n/a | n/a | n/a | n/a |
| 86 | 7510202 | 2310340H1 | SNP00004180 | 13 | 4124 | G | C | G | R1375 | n/a | n/a | n/a | n/a |
| 86 | 7510202 | 2846425H1 | SNP00049763 | 190 | 5108 | C | C | T | noncoding | n/d | n/a | n/a | n/a |
| 86 | 7510202 | 4619404H1 | SNP00004180 | 230 | 4129 | C | C | G | P1377 | n/a | n/a | n/a | n/a |
| 86 | 7510202 | 4619404H1 | SNP00024790 | 104 | 4003 | C | C | T | P1335 | n/d | n/a | n/a | n/a |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 86 | 7510202 | 5195240H1 | SNP00024790 | 123 | 4005 | C | C | T | P1335 | n/d | n/a | n/a | n/a |
| 86 | 7510202 | 5699760H1 | SNP00004180 | 24 | 4127 | G | C | G | G1376 | n/a | n/a | n/a | n/a |
| 86 | 7510202 | 7659234H1 | SNP00110738 | 261 | 4896 | C | A | C | noncoding | n/d | n/a | n/a | n/a |
| 87 | 7510203 | 5526052H2 | SNP00109646 | 120 | 994 | T | T | C | I54 | 0.74 | 0.66 | 0.87 | 0.83 |
| 87 | 7510203 | 6167684H1 | SNP00052166 | 128 | 2921 | T | C | T | noncoding | n/d | n/d | 1.00 | n/d |
| 87 | 7510203 | 6438120H1 | SNP00052167 | 442 | 3141 | T | T | G | noncoding | n/a | n/a | n/a | n/a |
| 87 | 7510203 | 6438174H1 | SNP00052167 | 447 | 3144 | T | T | G | noncoding | n/a | n/a | n/a | n/a |
| 87 | 7510203 | 6811774J1 | SNP00122473 | 96 | 562 | T | T | C | noncoding | 0.95 | 0.98 | 0.98 | 0.89 |
| 87 | 7510203 | 7604707H1 | SNP00122473 | 422 | 556 | T | T | C | noncoding | 0.95 | 0.98 | 0.98 | 0.89 |
| 88 | 7510208 | 1223476H1 | SNP00050176 | 208 | 7684 | C | C | T | noncoding | n/d | 0.87 | n/d | 0.98 |
| 88 | 7510208 | 1393432H1 | SNP00151799 | 53 | 3654 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 1421910H1 | SNP00116827 | 34 | 6310 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 1476551H1 | SNP00036803 | 115 | 3317 | T | T | C | noncoding | n/d | n/a | n/d | n/d |
| 88 | 7510208 | 1476551H1 | SNP00116828 | 89 | 3291 | G | G | A | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 1484827H1 | SNP00036802 | 92 | 5946 | T | T | C | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 1831572H1 | SNP00067230 | 96 | 7323 | A | A | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 1991459H1 | SNP00055209 | 13 | 5677 | T | T | C | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 1993878H1 | SNP00036803 | 23 | 6507 | T | T | C | noncoding | n/d | n/a | n/d | n/d |
| 88 | 7510208 | 2311751H1 | SNP00036803 | 203 | 3316 | T | T | C | noncoding | n/d | n/a | n/d | n/d |
| 88 | 7510208 | 2311751H1 | SNP00116828 | 177 | 3290 | G | G | A | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 2500080H1 | SNP00055209 | 124 | 2642 | T | T | C | A844 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 2572745H1 | SNP00151799 | 219 | 3653 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 2802374H1 | SNP00050176 | 105 | 7685 | C | C | T | noncoding | n/d | 0.87 | n/d | 0.98 |
| 88 | 7510208 | 2846083H1 | SNP00106013 | 228 | 328 | C | C | T | S73 | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 3087570H1 | SNP00036803 | 181 | 3297 | T | T | C | noncoding | n/d | n/a | n/d | n/d |
| 88 | 7510208 | 3087570H1 | SNP00116828 | 155 | 3271 | G | G | A | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 3398402H1 | SNP00106329 | 137 | 1697 | A | A | C | A529 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 3607395H1 | SNP00036803 | 16 | 6506 | T | T | C | noncoding | n/d | n/a | n/d | n/d |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 88 | 7510208 | 3744352H1 | SNP00067230 | 59 | 7322 | A | A | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 3753638H1 | SNP00116825 | 173 | 6859 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 3753638H1 | SNP00116826 | 270 | 6959 | C | C | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 3758446H1 | SNP00121526 | 193 | 4177 | G | G | T | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 3758446H1 | SNP00121527 | 89 | 4072 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 3822281H1 | SNP00151799 | 200 | 3652 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 3824109H1 | SNP00050176 | 146 | 7642 | T | C | T | noncoding | n/d | 0.87 | n/d | 0.98 |
| 88 | 7510208 | 3869114H1 | SNP00116827 | 50 | 3124 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 3941890H1 | SNP00106328 | 73 | 1539 | A | A | C | N477 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 3944290H1 | SNP00106329 | 230 | 1696 | A | A | C | E529 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 3950406H1 | SNP00062364 | 71 | 6893 | T | T | C | noncoding | 0.35 | 0.33 | 0.26 | 0.52 |
| 88 | 7510208 | 4082010H1 | SNP00050176 | 33 | 7683 | C | C | T | noncoding | n/d | 0.87 | n/d | 0.98 |
| 88 | 7510208 | 4093937H1 | SNP00116827 | 148 | 6307 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 4297233H1 | SNP00151799 | 120 | 3648 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 4342760H1 | SNP00067230 | 210 | 7321 | A | A | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 4456967H1 | SNP00055209 | 95 | 2643 | C | T | C | R845 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 4745011H1 | SNP00050176 | 193 | 7682 | C | C | T | noncoding | n/d | 0.87 | n/d | 0.98 |
| 88 | 7510208 | 4837070H1 | SNP00067230 | 84 | 7320 | A | A | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 5024420H1 | SNP00050176 | 201 | 7675 | C | C | T | noncoding | n/d | 0.87 | n/d | 0.98 |
| 88 | 7510208 | 5098681H2 | SNP00067230 | 108 | 7260 | A | A | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 5840056H1 | SNP00116825 | 179 | 3669 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 5840056H1 | SNP00116826 | 83 | 3765 | C | C | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 5971465H1 | SNP00036802 | 105 | 5944 | T | T | C | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 5972928H1 | SNP00050176 | 226 | 7681 | C | C | T | noncoding | n/d | 0.87 | n/d | 0.98 |
| 88 | 7510208 | 6253205H1 | SNP00106330 | 504 | 1945 | C | C | T | T612 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6253205H1 | SNP00106331 | 543 | 1984 | A | A | G | K625 | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 6436507H1 | SNP00067230 | 181 | 7316 | A | A | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 6572414H1 | SNP00106012 | 173 | 234 | C | C | T | L42 | n/a | n/a | n/a | n/a |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 88 | 7510208 | 6572451H1 | SNP00106014 | 553 | 584 | C | C | A | N158 | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 6574888H1 | SNP00060648 | 354 | 2717 | T | T | C | G869 | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 6618612H1 | SNP00106015 | 279 | 790 | A | A | G | E227 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 663140H1 | SNP00036802 | 174 | 5947 | T | T | C | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6757850J1 | SNP00060647 | 263 | 5464 | T | C | T | noncoding | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6762912J1 | SNP00093168 | 312 | 2382 | C | C | T | R758 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6763466J1 | SNP00060647 | 263 | 2427 | C | C | T | H773 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6765168H1 | SNP00106015 | 508 | 787 | A | A | G | E226 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6769705H1 | SNP00116826 | 424 | 3766 | C | C | G | noncoding | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 6814133H1 | SNP00106013 | 22 | 327 | C | C | T | P73 | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 6814133H1 | SNP00106014 | 282 | 583 | C | C | A | T158 | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 6872080H1 | SNP00106329 | 180 | 1653 | A | A | C | R515 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6872080H1 | SNP00106330 | 428 | 1904 | C | C | T | Y598 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6872080H1 | SNP00106331 | 467 | 1942 | A | A | G | Q611 | n/d | n/d | n/d | n/d |
| 88 | 7510208 | 6887907J1 | SNP00106015 | 109 | 1085 | G | A | G | G325 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6893301J1 | SNP00062364 | 361 | 6895 | T | T | C | noncoding | 0.35 | 0.33 | 0.26 | 0.52 |
| 88 | 7510208 | 6894642H1 | SNP00062364 | 441 | 3704 | T | T | C | noncoding | 0.35 | 0.33 | 0.26 | 0.52 |
| 88 | 7510208 | 6949007H1 | SNP00106015 | 524 | 999 | T | T | C | S297 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 6975081H1 | SNP00106012 | 152 | 233 | C | C | T | G41 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 7071727H1 | SNP00000461 | 353 | 7691 | T | T | C | noncoding | 0.47 | n/a | n/a | n/a |
| 88 | 7510208 | 7071727H1 | SNP00050176 | 401 | 7643 | C | C | T | noncoding | n/d | 0.87 | n/d | 0.98 |
| 88 | 7510208 | 7072747H1 | SNP00106015 | 307 | 788 | A | A | G | E226 | n/a | n/a | n/a | n/a |
| 88 | 7510208 | 7692966H2 | SNP00036803 | 28 | 6464 | A | A | G | noncoding | n/d | n/a | n/d | n/d |
| 88 | 7510208 | 7692966J2 | SNP00036803 | 265 | 6462 | A | A | G | noncoding | n/d | n/a | n/d | n/d |
| 90 | 7505294 | 1597548H1 | SNP00010943 | 166 | 419 | G | G | A | G129 | n/a | n/a | n/a | n/a |
| 90 | 7505294 | 1915171H1 | SNP00041595 | 37 | 1002 | T | T | C | W324 | n/d | n/a | n/a | n/a |
| 90 | 7505294 | 2104778H1 | SNP00115560 | 95 | 95 | C | C | T | P21 | n/d | n/d | n/a | n/a |
| 90 | 7505294 | 2615711H1 | SNP00041595 | 215 | 1003 | T | T | C | L324 | n/d | n/a | n/a | n/a |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 90 | 7505294 | 2679723H1 | SNP00010943 | 24 | 418 | G | G | A | G129 | n/a | n/a | n/a | n/a |
| 90 | 7505294 | 3763389H1 | SNP00041595 | 6 | 1001 | T | T | C | P323 | n/d | n/a | n/a | n/a |
| 90 | 7505294 | 3781729H1 | SNP00010943 | 123 | 417 | G | G | A | G129 | n/a | n/a | n/a | n/a |
| 90 | 7505294 | 3941309H1 | SNP00041595 | 246 | 996 | T | T | C | S322 | n/d | n/a | n/a | n/a |
| 90 | 7505294 | 5848280H1 | SNP00041595 | 30 | 997 | T | T | C | L322 | n/d | n/a | n/a | n/a |
| 90 | 7505294 | 6147146H1 | SNP00010943 | 295 | 413 | G | G | A | R127 | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 1355009H1 | SNP00025254 | 19 | 2357 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 1355009H1 | SNP00025255 | 95 | 2433 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 1367924H1 | SNP00004462 | 113 | 3354 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 1644485H1 | SNP00025253 | 32 | 1603 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 2047058H1 | SNP00025254 | 82 | 2356 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 2047058H1 | SNP00025255 | 6 | 2432 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 2325919H1 | SNP00004462 | 67 | 3351 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 2824251H1 | SNP00004462 | 114 | 3352 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 3495656H1 | SNP00025254 | 159 | 2354 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 3495656H1 | SNP00025255 | 235 | 2430 | G | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 4329080H1 | SNP00025254 | 65 | 2355 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 4329080H1 | SNP00025255 | 141 | 2431 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 4525326H1 | SNP00025254 | 200 | 2352 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 4647714H1 | SNP00025253 | 196 | 1599 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 4980914H1 | SNP00025253 | 55 | 1601 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 91 | 7505631 | 5681630H1 | SNP00025253 | 198 | 1602 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 92 | 7506561 | 1667982H1 | SNP00071142 | 118 | 316 | C | C | A | R26 | 1.00 | n/a | n/a | n/a |
| 93 | 7510733 | 1218582H1 | SNP00052605 | 135 | 2689 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 1222021H1 | SNP00040592 | 51 | 2412 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 1222329H1 | SNP00040592 | 107 | 2419 | A | C | A | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 1559060H1 | SNP00040591 | 61 | 2107 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 1559060H1 | SNP00047819 | 83 | 2129 | G | C | G | noncoding | 0.06 | 0.18 | 0.09 | 0.06 |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 93 | 7510733 | 3693413H1 | SNP00040592 | 227 | 2415 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 3840578H1 | SNP00040592 | 102 | 2416 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 3840668H1 | SNP00040591 | 125 | 2104 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 3840668H1 | SNP00047819 | 147 | 2126 | G | C | G | noncoding | 0.06 | 0.18 | 0.09 | 0.06 |
| 93 | 7510733 | 4410720H1 | SNP00052604 | 19 | 903 | G | G | A | A233 | 0.89 | 0.82 | 0.87 | 0.91 |
| 93 | 7510733 | 4414082H1 | SNP00040591 | 59 | 2105 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 4414082H1 | SNP00047819 | 81 | 2127 | G | C | G | noncoding | 0.06 | 0.18 | 0.09 | 0.06 |
| 93 | 7510733 | 4414445H1 | SNP00052605 | 163 | 2688 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 478403H1 | SNP00052605 | 87 | 2690 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 557755H1 | SNP00040592 | 133 | 2417 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 559550H1 | SNP00040591 | 60 | 2106 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 93 | 7510733 | 559550H1 | SNP00047819 | 82 | 2128 | C | C | G | noncoding | 0.06 | 0.18 | 0.09 | 0.06 |
| 93 | 7510733 | 5824493H1 | SNP00052604 | 355 | 899 | A | G | A | L231 | 0.89 | 0.82 | 0.87 | 0.91 |
| 93 | 7510733 | 6896821H1 | SNP00052603 | 140 | 205 | A | A | G | noncoding | 0.95 | n/d | n/d | 0.97 |
| 94 | 7510734 | 1218582H1 | SNP00052605 | 135 | 2774 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 1222021H1 | SNP00040592 | 51 | 2497 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 1222329H1 | SNP00040592 | 107 | 2504 | A | C | A | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 1559060H1 | SNP00040591 | 61 | 2192 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 1559060H1 | SNP00047819 | 83 | 2214 | G | C | G | noncoding | 0.06 | 0.18 | 0.09 | 0.06 |
| 94 | 7510734 | 3693413H1 | SNP00040592 | 227 | 2500 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 3840578H1 | SNP00040592 | 102 | 2501 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 3840668H1 | SNP00040591 | 125 | 2189 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 3840668H1 | SNP00047819 | 147 | 2211 | G | C | G | noncoding | 0.06 | 0.18 | 0.09 | 0.06 |
| 94 | 7510734 | 4410720H1 | SNP00052604 | 19 | 988 | G | G | A | noncoding | 0.89 | 0.82 | 0.87 | 0.91 |
| 94 | 7510734 | 4414082H1 | SNP00040591 | 59 | 2190 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 4414082H1 | SNP00047819 | 81 | 2212 | G | C | G | noncoding | 0.06 | 0.18 | 0.09 | 0.06 |
| 94 | 7510734 | 4414445H1 | SNP00052605 | 163 | 2773 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 478403H1 | SNP00052605 | 87 | 2775 | C | C | T | noncoding | n/a | n/a | n/a | n/a |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 94 | 7510734 | 557755H1 | SNP00040592 | 133 | 2502 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 559550H1 | SNP00040591 | 60 | 2191 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 94 | 7510734 | 559550H1 | SNP00047819 | 82 | 2213 | C | C | G | noncoding | 0.06 | 0.18 | 0.09 | 0.06 |
| 94 | 7510734 | 5824493H1 | SNP00052604 | 355 | 984 | A | G | A | noncoding | 0.89 | 0.82 | 0.87 | 0.91 |
| 94 | 7510734 | 6896821H1 | SNP00052603 | 140 | 205 | A | A | G | noncoding | 0.95 | n/d | n/d | 0.97 |
| 95 | 7503977 | 1627645H1 | SNP00033082 | 5 | 1937 | C | C | T | noncoding | n/d | n/a | n/a | n/a |
| 95 | 7503977 | 3027777F6 | SNP00033081 | 49 | 1624 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 95 | 7503977 | 6800356J1 | SNP00107876 | 418 | 711 | G | G | T | L173 | n/d | n/a | n/a | n/a |
| 96 | 7505084 | 2707270F6 | SNP00152729 | 184 | 1603 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 97 | 7506950 | 1412498H1 | SNP00112954 | 200 | 439 | G | G | A | K132 | n/a | n/a | n/a | n/a |
| 98 | 7506951 | 1412498H1 | SNP00112954 | 200 | 439 | G | G | A | K132 | n/a | n/a | n/a | n/a |
| 100 | 7506956 | 1412498H1 | SNP00112954 | 200 | 439 | G | G | A | K132 | n/a | n/a | n/a | n/a |
| 101 | 7506959 | 1412498H1 | SNP00112954 | 200 | 439 | G | G | A | K132 | n/a | n/a | n/a | n/a |
| 102 | 7506960 | 7233773H1 | SNP00112954 | 137 | 448 | A | G | A | noncoding | n/a | n/a | n/a | n/a |
| 103 | 7510540 | 2923154F6 | SNP00019786 | 209 | 1571 | A | A | C | noncoding | 0.07 | n/a | n/a | n/a |
| 103 | 7510540 | 6930765H1 | SNP00098509 | 28 | 356 | A | G | A | noncoding | n/a | n/a | n/a | n/a |
| 104 | 7510545 | 1275854F6 | SNP00124648 | 18 | 175 | G | G | A | V39 | n/d | n/d | n/d | n/d |
| 104 | 7510545 | 2347746H1 | SNP00041565 | 38 | 787 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 104 | 7510545 | 2347746H1 | SNP00041566 | 144 | 893 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 104 | 7510545 | 7276247H2 | SNP00124648 | 103 | 173 | G | G | A | G38 | n/d | n/d | n/d | n/d |
| 104 | 7510545 | 7602268J1 | SNP00124648 | 187 | 164 | G | G | A | G35 | n/d | n/d | n/d | n/d |
| 104 | 7510545 | 7741944J1 | SNP00124648 | 555 | 153 | G | G | A | E31 | n/d | n/d | n/d | n/d |
| 105 | 7510654 | 1250172H1 | SNP00098839 | 202 | 1274 | G | G | A | noncoding | n/a | n/a | n/a | n/a |
| 105 | 7510654 | 1416107F6 | SNP00007052 | 479 | 1507 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 105 | 7510654 | 1416107T6 | SNP00007052 | 153 | 1582 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 105 | 7510654 | 1416107T6 | SNP00032083 | 118 | 1617 | G | G | T | noncoding | n/a | n/a | n/a | n/a |
| 105 | 7510654 | 1553708H1 | SNP00007052 | 61 | 1506 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 105 | 7510654 | 1553708H1 | SNP00032083 | 26 | 1541 | G | G | T | noncoding | n/a | n/a | n/a | n/a |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 105 | 7510654 | 1750720F6 | SNP00001888 | 154 | 826 | C | C | T | G260 | n/a | n/a | n/a | n/a |
| 105 | 7510654 | 1750720F6 | SNP00074165 | 198 | 870 | T | T | G | V275 | n/a | n/a | n/a | n/a |
| 105 | 7510654 | 2185847F6 | SNP00001888 | 164 | 827 | T | C | T | F261 | n/a | n/a | n/a | n/a |
| 105 | 7510654 | 2185847F6 | SNP00074165 | 208 | 871 | T | T | G | V275 | n/a | n/a | n/a | n/a |
| 106 | 7510660 | 1208437R1 | SNP00076070 | 188 | 2807 | C | C | T | S893 | n/a | n/a | n/a | n/a |
| 106 | 7510660 | 2723676F6 | SNP00116349 | 246 | 1935 | C | C | T | P603 | n/a | n/a | n/a | n/a |
| 106 | 7510660 | 3392285H1 | SNP00116349 | 209 | 1932 | C | C | T | L602 | n/a | n/a | n/a | n/a |
| 106 | 7510660 | 5401847F6 | SNP00076069 | 316 | 1721 | C | C | T | C531 | n/d | n/a | n/a | n/a |
| 107 | 7510661 | 1208437R1 | SNP00076070 | 188 | 2737 | C | C | T | noncoding | n/a | n/a | n/a | n/a |
| 107 | 7510661 | 2723676F6 | SNP00116349 | 246 | 1935 | C | C | T | P603 | n/a | n/a | n/a | n/a |
| 107 | 7510661 | 3392285H1 | SNP00116349 | 209 | 1932 | C | C | T | L602 | n/a | n/a | n/a | n/a |
| 107 | 7510661 | 5401847F6 | SNP00076069 | 316 | 1721 | C | C | T | C531 | n/d | n/a | n/a | n/a |
| 108 | 7510680 | 1443748R1 | SNP00149102 | 53 | 1909 | G | G | A | noncoding | n/a | n/a | n/a | n/a |
| 108 | 7510680 | 1443748T6 | SNP00149102 | 5 | 1910 | G | G | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 1288322H1 | SNP00020995 | 117 | 1419 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 1307614H1 | SNP00020993 | 11 | 466 | C | C | T | T129 | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 1954824H1 | SNP00020995 | 3 | 1416 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 2445108H1 | SNP00020995 | 66 | 1415 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 2701206H1 | SNP00020995 | 48 | 1418 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 2750471H1 | SNP00020994 | 46 | 1358 | G | G | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 3508476H1 | SNP00020992 | 20 | 110 | C | G | C | T10 | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 4375677H1 | SNP00020993 | 76 | 465 | C | C | T | P129 | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 4595210H1 | SNP00020993 | 108 | 464 | C | C | T | A128 | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 4649323H1 | SNP00020993 | 247 | 467 | C | C | T | T129 | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 4764638H1 | SNP00020994 | 210 | 1356 | G | G | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 5100029H1 | SNP00020995 | 225 | 1421 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 5951826H1 | SNP00020995 | 214 | 1410 | C | C | A | noncoding | n/a | n/a | n/a | n/a |
| 109 | 7505145 | 6846420H1 | SNP00020994 | 194 | 1361 | G | G | A | noncoding | n/a | n/a | n/a | n/a |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 109 | 7505145 | 7344481H1 | SNP00094043 | 138 | 1067 | G | C | G | G329 | n/d | n/a | n/a | n/a |
| 110 | 7505162 | 139886H1 | SNP00036849 | 97 | 1345 | A | G | A | noncoding | 0.77 | n/a | n/a | n/a |
| 110 | 7505162 | 1552730H1 | SNP00075257 | 80 | 681 | C | C | T | A156 | n/a | n/a | n/a | n/a |
| 110 | 7505162 | 1620127H1 | SNP00009022 | 98 | 1542 | G | G | C | noncoding | n/a | n/a | n/a | n/a |
| 110 | 7505162 | 1620504H1 | SNP00009022 | 99 | 1543 | G | G | C | noncoding | n/a | n/a | n/a | n/a |
| 110 | 7505162 | 2679787H1 | SNP00036849 | 211 | 1349 | G | G | A | noncoding | 0.77 | n/a | n/a | n/a |
| 110 | 7505162 | 3556418H1 | SNP00036849 | 174 | 1352 | G | G | A | noncoding | 0.77 | n/a | n/a | n/a |
| 110 | 7505162 | 3852765H1 | SNP00009022 | 121 | 1541 | G | G | C | noncoding | n/a | n/a | n/a | n/a |
| 110 | 7505162 | 6746110H1 | SNP00009022 | 554 | 1538 | G | G | C | noncoding | n/a | n/a | n/a | n/a |
| 111 | 7505469 | 217011H1 | SNP00011018 | 91 | 2065 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 111 | 7505469 | 2402129H1 | SNP00011018 | 80 | 2066 | G | A | G | noncoding | n/a | n/a | n/a | n/a |
| 112 | 7505475 | 1307948H1 | SNP00023319 | 77 | 4003 | A | G | A | noncoding | 0.57 | 0.28 | 0.36 | 0.65 |
| 112 | 7505475 | 1307948H1 | SNP00023320 | 201 | 4127 | C | C | T | noncoding | n/d | n/d | n/d | n/d |
| 112 | 7505475 | 1307948H1 | SNP00072137 | 91 | 4017 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 112 | 7505475 | 1965082H1 | SNP00023319 | 146 | 4002 | G | G | A | noncoding | 0.57 | 0.28 | 0.36 | 0.65 |
| 112 | 7505475 | 1965082H1 | SNP00072137 | 160 | 4016 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 112 | 7505475 | 2764315H1 | SNP00026116 | 169 | 3871 | T | T | C | noncoding | n/d | n/a | n/a | n/a |
| 112 | 7505475 | 3781324H1 | SNP00026116 | 174 | 3868 | C | T | C | noncoding | n/d | n/a | n/a | n/a |
| 112 | 7505475 | 3781324H1 | SNP00139248 | 272 | 3966 | T | C | T | noncoding | n/a | n/a | n/a | n/a |
| 112 | 7505475 | 3948943H1 | SNP00023320 | 89 | 4126 | C | C | T | noncoding | n/d | n/d | n/d | n/d |
| 112 | 7505475 | 5704531H1 | SNP00023320 | 113 | 4124 | T | C | T | noncoding | n/d | n/d | n/d | n/d |
| 112 | 7505475 | 5704531H1 | SNP00072137 | 3 | 4014 | A | A | G | noncoding | n/a | n/a | n/a | n/a |
| 113 | 7505568 | 2731808H1 | SNP00035633 | 100 | 129 | G | A | G | G25 | 0.73 | 0.30 | 0.74 | 0.73 |
| 113 | 7505568 | 4552729H1 | SNP00035634 | 167 | 460 | T | C | T | noncoding | n/a | n/a | n/a | n/a |
| 116 | 7510541 | 2514486H1 | SNP00142846 | 30 | 97 | C | C | G | noncoding | n/a | n/a | n/a | n/a |
| 116 | 7510541 | 574452H1 | SNP00142846 | 60 | 96 | C | C | G | noncoding | n/a | n/a | n/a | n/a |
| 116 | 7510541 | 5758634H1 | SNP00142846 | 45 | 94 | C | C | G | noncoding | n/a | n/a | n/a | n/a |
| 117 | 7510923 | 2514486H1 | SNP00142846 | 30 | 97 | C | C | G | noncoding | n/a | n/a | n/a | n/a |

# Table 8

| SEQ ID NO: | PID | EST ID | SNP ID | EST SNP | CB1 SNP | EST Allele | Allele 1 | Allele 2 | Amino Acid | Caucasian Allele 1 frequency | African Allele 1 frequency | Asian Allele 1 frequency | Hispanic Allele 1 frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 117 | 7510923 | 574452H1 | SNP00142846 | 60 | 96 | C | C | G | noncoding | n/a | n/a | n/a | n/a |
| 117 | 7510923 | 5758634H1 | SNP00142846 | 45 | 94 | C | C | G | noncoding | n/a | n/a | n/a | n/a |
| 118 | 7510984 | 1270543H1 | SNP00121815 | 232 | 3202 | G | G | A | R1049 | n/a | n/a | n/a | n/a |
| 118 | 7510984 | 2402461H1 | SNP00051864 | 90 | 4465 | C | C | T | noncoding | n/d | n/a | n/a | n/a |
| 118 | 7510984 | 6559367H1 | SNP00051863 | 151 | 3877 | G | G | A | R1274 | 0.50 | 0.18 | 0.74 | 0.38 |
| 118 | 7510984 | 6908670H1 | SNP00121813 | 251 | 262 | C | C | T | P69 | n/a | n/a | n/a | n/a |
| 118 | 7510984 | 6908670H1 | SNP00121814 | 374 | 385 | C | C | T | A110 | n/d | 0.95 | n/d | n/d |

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59,

b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:10-11, SEQ ID NO:20-21, SEQ ID NO:23, SEQ ID NO:25-29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:38-40, SEQ ID NO:43-45, SEQ ID NO:49-51, and SEQ ID NO:56,

c) a polypeptide comprising a naturally occurring amino acid sequence at least 91% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:3 and SEQ ID NO:54,

d) a polypeptide comprising a naturally occurring amino acid sequence at least 92% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:19 and SEQ ID NO:22,

e) a polypeptide comprising a naturally occurring amino acid sequence at least 93% identical to the amino acid sequence of SEQ ID NO:46,

f) a polypeptide comprising a naturally occurring amino acid sequence at least 94% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:14-15,

g) a polypeptide comprising a naturally occurring amino acid sequence at least 95% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:13, SEQ ID NO:18, SEQ ID NO:36-37 and SEQ ID NO:52,

h) a polypeptide comprising a naturally occurring amino acid sequence at least 96% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:16,

i) a polypeptide comprising a naturally occurring amino acid sequence at least 97% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:30, SEQ ID NO:32 and SEQ ID NO:57-58,

j) a polypeptide comprising a naturally occurring amino acid sequence at least 98% identical to the amino acid sequence of SEQ ID NO:59,

k) a polypeptide comprising a naturally occurring amino acid sequence at least 99% identical to the amino acid sequence of SEQ ID NO:12,

l) a polypeptide consisting essentially of a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:34, SEQ ID NO:41-42, SEQ ID NO:47, and SEQ ID NO:55.

m) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, and

n) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method of producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant

polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:60-118,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:60, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:73, SEQ ID NO:76, SEQ ID NO:78-79, SEQ ID NO:81, SEQ ID NO:83-85, SEQ ID NO:90-92, SEQ ID NO:108 and SEQ ID NO:113,

c) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 91% identical to the polynucleotide sequence of SEQ ID NO:70,

d) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 92% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:71 and SEQ ID NO:87,

e) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 93% identical to the polynucleotide sequence of SEQ ID NO:115,

f) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 94% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:68 and SEQ ID NO:117,

g) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 95% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:63, SEQ ID NO:66 and SEQ ID NO:118,

h) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 96% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:64 and SEQ ID NO:75,

i) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 97% identical to the polynucleotide sequence of SEQ ID NO:106,

j) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 98% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:80, SEQ ID NO:88, SEQ ID NO:95, SEQ ID NO:109 and SEQ ID NO:116,

k) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 99% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:62, SEQ ID NO:82, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:96, SEQ ID NO:99 and SEQ ID NO:104,

l) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:61, SEQ ID NO:69, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:93-94, SEQ ID NO:97-98, SEQ ID NO:100-103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:110-112 and SEQ ID NO:114,

m) a polynucleotide complementary to a polynucleotide of a),

n) a polynucleotide complementary to a polynucleotide of b),

o) a polynucleotide complementary to a polynucleotide of c),

p) a polynucleotide complementary to a polynucleotide of d),

q) a polynucleotide complementary to a polynucleotide of e),

r) a polynucleotide complementary to a polynucleotide of f),

s) a polynucleotide complementary to a polynucleotide of g),

t) a polynucleotide complementary to a polynucleotide of h),

u) a polynucleotide complementary to a polynucleotide of i),

v) a polynucleotide complementary to a polynucleotide of j),

w) a polynucleotide complementary to a polynucleotide of k),

x) a polynucleotide complementary to a polynucleotide of l), and

y) an RNA equivalent of a)-x).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

19. A method for treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional TRICH, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound; the method comprising:

a) treating a biological sample containing nucleic acids with the test compound,

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

c) quantifying the amount of hybridization complex, and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A method for a diagnostic test for a condition or disease associated with the expression of TRICH in a biological sample, the method comprising:

a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of TRICH in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, further comprising a label.

35. A method of diagnosing a condition or disease associated with the expression of TRICH in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibodies from the animal, and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-59, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibody producing cells from the animal,

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,

d) culturing the hybridoma cells, and

e) isolating from the culture monoclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59 in a sample, the method comprising:

a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59 from a sample, the method comprising:

a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-59.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

92. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

93. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

94. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

95. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

96. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

97. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

98. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.

99. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.

100. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:45.

101. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:46.

102. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:47.

103. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:48.

104. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:49.

105. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:50.

106. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:51.

107. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:52.

108. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:53.

109. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:54.

110. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:55.

111. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:56.

112. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:57.

113. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:58.

114. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:59.

115. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:60.

116. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:61.

117. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:62.

118. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:63.

119. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:64.

120. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:65.

121. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:66.

122. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:67.

123. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:68.

124. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:69.

125. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:70.

126. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:71.

127. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:72.

128. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:73.

129. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:74.

130. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:75.

131. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:76.

132. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:77.

133. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:78.

134. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:79.

135. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:80.

136. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:81.

137. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:82.

138. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:83.

139. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:84.

140. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:85.

141. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:86.

142. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:87.

143. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:88.

144. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:89.

145. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:90.

146. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:91.

147. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:92.

148. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:93.

149. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:94.

150. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:95.

151. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:96.

152. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:97.

153. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:98.

154. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:99.

155. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:100.

156. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:101.

157. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:102.

158. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:103.

159. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:104.

160. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:105.

161. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:106.

162. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:107.

163. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:108.

164. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:109.

165. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:110.

166. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:111.

167. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:112.

168. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:113.

169. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:114.

170. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:115.

171. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:116.

172. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:117.

173. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:118.

```
<110> INCYTE CORPORATION
      MARQUIS, Joseph P.
      LEE, Soo Y.
      EMERLING, Brooke M.
      HAFALLIA, April J.A.
      KHARE, Reena
      KABLE, Amy E.
      RICHARDSON, Thomas W.
      SWARNAKAR, Anita
      CHAWLA, Narinder K.
      BECHA, Shanya D.
      MASON, Patricia M.
      ELLIOTT, Vicki S.
      RAMKUMAR, Jayalaxmi
      GRIFFIN, Jennifer A.
      TRAN, Uyen K.
      ISON, Craig H.
      LINDQUIST, Erika A.
      JIANG, Xin
      JACKSON, Alan A.
      WILSON, Amy D.
      JIN, Pei
      CHANG, Hsin-Ru

<120> TRANSPORTERS AND ION CHANNELS

<130> PF-1397 PCT

<140> To Be Assigned
<141> Herewith

<150> US 60/368,840
<151> 2002-03-28

<150> US 60/375,637
<151> 2002-04-26

<160> 118
<170> PERL Program

<210> 1
<211> 195
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509332CD1

<400> 1
Met Asp Met Ala Trp Gln Met Met Gln Leu Leu Leu Leu Ala Leu
  1               5                  10                  15
Val Thr Ala Ala Gly Ser Ala Gln Pro Arg Ser Ala Arg Ala Arg
                 20                  25                  30
Thr Asp Leu Leu Asn Val Cys Met Asn Ala Lys His His Lys Thr
                 35                  40                  45
```

Gln Pro Ser Pro Glu Asp Glu Leu Tyr Gly Gln Cys Ser Pro Trp
50 55 60
Lys Lys Asn Ala Cys Cys Thr Ala Ser Thr Ser Gln Glu Leu His
65 70 75
Lys Asp Thr Ser Arg Leu Tyr Asn Phe Asn Trp Asp His Cys Glu
80 85 90
Arg Trp Trp Glu Asp Cys Arg Thr Ser Tyr Thr Cys Lys Ser Asn
95 100 105
Trp His Lys Gly Trp Asn Trp Thr Ser Gly Ile Asn Glu Cys Pro
110 115 120
Ala Gly Ala Leu Cys Ser Thr Phe Glu Ser Tyr Phe Pro Thr Pro
125 130 135
Ala Ala Leu Cys Glu Gly Leu Trp Ser His Ser Phe Lys Val Ser
140 145 150
Asn Tyr Ser Arg Gly Ser Gly Arg Cys Ile Gln Met Trp Phe Asp
155 160 165
Ser Ala Gln Gly Asn Pro Asn Glu Glu Val Ala Lys Phe Tyr Ala
170 175 180
Ala Ala Met Asn Ala Gly Ala Pro Ser Arg Gly Ile Ile Asp Ser
185 190 195

<210> 2
<211> 138
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509102CD1

<400> 2
Met Gly Pro Ser Cys Pro Val Phe Leu Ser Phe Thr Lys Leu Gly
1 5 10 15
Leu Trp Trp Leu Leu Leu Thr Pro Ala Gly Gly Glu Glu Ala Lys
20 25 30
Arg Pro Pro Pro Arg Ala Pro Gly Asp Pro Leu Ser Ser Pro Ser
35 40 45
Pro Thr Ala Leu Pro Gln Gly Gly Ser His Thr Glu Thr Glu Asp
50 55 60
Arg Leu Phe Lys His Leu Phe Arg Gly Tyr Asn Arg Trp Ala Arg
65 70 75
Pro Val Pro Asn Thr Ser Asp Val Asp Glu Lys Asn Gln Met Met
80 85 90
Thr Thr Asn Val Trp Leu Lys Gln Glu Trp Ser Asp Tyr Lys Leu
95 100 105
Arg Trp Asn Pro Thr Asp Phe Gly Asn Ile Thr Ser Leu Arg Val
110 115 120
Pro Ser Glu Met Ile Trp Ile Pro Asp Ile Val Leu Tyr Asn Lys
125 130 135
Thr Ala Arg

<210> 3
<211> 355
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509132CD1

<400> 3

```
Met Ser Trp Arg Cys Trp Gly Ala Ala Ser Trp Ala Trp Pro Met
  1               5                  10                  15
Leu Leu Pro Pro Met Lys Cys Ser Ser Leu Asp Asp Ser Ser Leu
                 20                  25                  30
Ala Pro Thr Gln Val Leu Gly Leu Glu Ser Leu Leu Gly Thr Ala
                 35                  40                  45
Ser Leu Trp Pro Leu Leu Leu Gly Leu Thr Val Leu Pro Ala Leu
                 50                  55                  60
Leu Gln Leu Val Leu Leu Pro Phe Cys Pro Glu Ser Pro Arg Tyr
                 65                  70                  75
Leu Tyr Ile Ile Gln Asn Leu Glu Gly Pro Ala Arg Lys Ser Leu
                 80                  85                  90
Lys Arg Leu Thr Gly Trp Ala Asp Val Ser Gly Val Leu Ala Glu
                 95                 100                 105
Leu Lys Asp Glu Lys Arg Lys Leu Glu Arg Glu Arg Pro Leu Ser
                110                 115                 120
Leu Leu Gln Leu Leu Gly Ser Arg Thr His Arg Gln Pro Leu Ile
                125                 130                 135
Ile Ala Val Val Leu Gln Leu Ser Gln Gln Leu Ser Gly Ile Asn
                140                 145                 150
Ala Val Phe Tyr Tyr Ser Thr Ser Ile Phe Glu Thr Ala Gly Val
                155                 160                 165
Gly Gln Pro Ala Tyr Ala Thr Ile Gly Ala Gly Val Val Asn Thr
                170                 175                 180
Val Phe Thr Leu Val Ser Val Leu Leu Val Glu Arg Ala Gly Arg
                185                 190                 195
Arg Thr Leu His Leu Leu Gly Leu Ala Gly Met Cys Gly Cys Ala
                200                 205                 210
Ile Leu Met Thr Val Ala Leu Leu Leu Leu Glu Arg Val Pro Ala
                215                 220                 225
Met Ser Tyr Val Ser Ile Val Ala Ile Phe Gly Phe Val Ala Phe
                230                 235                 240
Phe Glu Ile Gly Pro Gly Pro Ile Pro Trp Phe Ile Val Ala Glu
                245                 250                 255
Leu Phe Ser Gln Gly Pro Arg Pro Ala Ala Met Ala Val Ala Gly
                260                 265                 270
Phe Ser Asn Trp Thr Ser Asn Phe Ile Ile Gly Met Gly Phe Gln
                275                 280                 285
Tyr Val Ala Glu Ala Met Gly Pro Tyr Val Phe Leu Leu Phe Ala
                290                 295                 300
Val Leu Leu Leu Gly Phe Phe Ile Phe Thr Phe Leu Arg Val Pro
                305                 310                 315
Glu Thr Arg Gly Arg Thr Phe Asp Gln Ile Ser Ala Ala Phe His
                320                 325                 330
Arg Thr Pro Ser Leu Leu Glu Gln Glu Val Lys Pro Ser Thr Glu
                335                 340                 345
Leu Glu Tyr Leu Gly Pro Asp Glu Asn Asp
                350                 355
```

<210> 4
<211> 380
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509136CD1

<400> 4
Met Ser Thr Lys Val Tyr Leu Asp Leu Glu Trp Thr Asp Tyr Arg
1 5 10 15
Leu Ser Trp Asp Pro Ala Glu His Asp Gly Ile Asp Ser Leu Arg
20 25 30
Ile Thr Ala Glu Ser Val Trp Leu Pro Asp Val Val Leu Leu Asn
35 40 45
Asn Asn Asp Gly Asn Phe Asp Val Ala Leu Asp Ile Ser Val Val
50 55 60
Val Ser Ser Asp Gly Ser Val Arg Trp Gln Pro Pro Gly Ile Tyr
65 70 75
Arg Ser Ser Cys Ser Ile Gln Val Thr Tyr Phe Pro Phe Asp Trp
80 85 90
Gln Asn Cys Thr Met Val Phe Ser Ser Tyr Ser Tyr Asp Ser Ser
95 100 105
Glu Val Ser Leu Gln Thr Gly Leu Gly Pro Asp Gly Gln Gly His
110 115 120
Gln Glu Ile His Ile His Glu Gly Thr Phe Ile Glu Asn Gly Gln
125 130 135
Trp Glu Ile Ile His Lys Pro Ser Arg Leu Ile Gln Pro Pro Gly
140 145 150
Asp Pro Arg Gly Gly Arg Glu Gly Gln Arg Gln Glu Val Ile Phe
155 160 165
Tyr Leu Ile Ile Arg Arg Lys Pro Leu Phe Tyr Leu Val Asn Val
170 175 180
Ile Ala Pro Cys Ile Leu Ile Thr Leu Leu Ala Ile Phe Val Phe
185 190 195
Tyr Leu Pro Pro Asp Ala Val Ile Leu Ser Val Val Val Leu Asn
200 205 210
Leu His His Arg Ser Pro His Thr His Gln Met Pro Leu Trp Val
215 220 225
Arg Gln Ile Phe Ile His Lys Leu Pro Leu Tyr Leu Arg Leu Lys
230 235 240
Arg Pro Lys Pro Glu Arg Asp Leu Met Pro Glu Pro Pro His Cys
245 250 255
Ser Ser Pro Gly Ser Gly Trp Gly Arg Gly Thr Asp Glu Tyr Phe
260 265 270
Ile Arg Lys Pro Pro Ser Asp Phe Leu Phe Pro Lys Pro Asn Arg
275 280 285
Phe Gln Pro Glu Leu Ser Ala Pro Asp Leu Arg Arg Phe Ile Asp
290 295 300
Gly Pro Asn Arg Ala Val Ala Leu Leu Pro Glu Leu Arg Glu Val
305 310 315
Val Ser Ser Ile Ser Tyr Ile Ala Arg Gln Leu Gln Glu Gln Glu
320 325 330
Asp His Asp Ala Leu Lys Glu Asp Trp Gln Phe Val Ala Met Val
335 340 345

Val Asp Arg Leu Phe Leu Trp Thr Phe Ile Ile Phe Thr Ser Val
350 355 360
Gly Thr Leu Val Ile Phe Leu Asp Ala Thr Tyr His Leu Pro Pro
365 370 375
Pro Asp Pro Phe Pro
380

<210> 5
<211> 375
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509178CD1

<400> 5
Met Glu Pro Trp Pro Leu Leu Leu Leu Phe Ser Leu Cys Ser Ala
1 5 10 15
Gly Leu Val Leu Gly Ser Glu His Glu Thr Arg Leu Val Ala Lys
20 25 30
Leu Phe Lys Asp Tyr Ser Ser Val Val Arg Pro Val Glu Asp His
35 40 45
Arg Gln Val Val Glu Val Thr Val Gly Leu Gln Leu Ile Gln Leu
50 55 60
Ile Asn Val Asp Glu Val Asn Gln Ile Val Thr Thr Asn Val Arg
65 70 75
Leu Lys Gln Asn Cys Ser Met Lys Leu Gly Thr Trp Thr Tyr Asp
80 85 90
Gly Ser Val Val Ala Ile Asn Pro Glu Ser Asp Gln Pro Asp Leu
95 100 105
Ser Asn Phe Met Glu Ser Gly Glu Trp Val Ile Lys Glu Ser Arg
110 115 120
Gly Trp Lys His Ser Val Thr Tyr Ser Cys Cys Pro Asp Thr Pro
125 130 135
Tyr Leu Asp Ile Thr Tyr His Phe Val Met Gln Arg Leu Pro Leu
140 145 150
Tyr Phe Ile Val Asn Val Ile Ile Pro Cys Leu Leu Phe Ser Phe
155 160 165
Leu Thr Gly Leu Val Phe Tyr Leu Pro Thr Asp Ser Gly Glu Lys
170 175 180
Met Thr Leu Ser Ile Ser Val Leu Leu Ser Leu Thr Val Phe Leu
185 190 195
Leu Val Ile Val Glu Leu Ile Pro Ser Thr Ser Ser Ala Val Pro
200 205 210
Leu Ile Gly Lys Tyr Met Leu Phe Thr Met Val Phe Val Ile Ala
215 220 225
Ser Ile Ile Ile Thr Val Ile Val Ile Asn Thr His His Arg Ser
230 235 240
Pro Ser Thr His Val Met Pro Asn Trp Val Arg Lys Val Phe Ile
245 250 255
Asp Thr Ile Pro Asn Ile Met Phe Phe Ser Thr Met Lys Arg Pro
260 265 270
Ser Arg Glu Lys Gln Asp Lys Lys Ile Phe Thr Glu Asp Ile Asp
275 280 285
Ile Ser Asp Ile Ser Gly Lys Pro Gly Pro Pro Pro Met Gly Phe

290 295 300
His Ser Pro Leu Ile Lys His Pro Glu Val Lys Ser Ala Ile Glu
305 310 315
Gly Ile Lys Tyr Ile Ala Glu Thr Met Lys Ser Asp Gln Glu Ser
320 325 330
Asn Asn Ala Ala Ala Glu Trp Lys Tyr Val Ala Met Val Met Asp
335 340 345
His Ile Leu Leu Gly Val Phe Met Leu Val Cys Ile Ile Gly Thr
350 355 360
Leu Ala Val Phe Ala Gly Arg Leu Ile Glu Leu Asn Gln Gln Gly
365 370 375

<210> 6
<211> 153
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509214CD1

<400> 6
Met Ala Pro Pro Trp Val Pro Ala Met Gly Phe Thr Leu Ala Pro
1 5 10 15
Ser His Gly Val Arg Leu Leu Pro Gly Leu Glu Arg Ala Gly Arg
20 25 30
Leu His Arg Glu Gly Cys Gly Ser Pro Gly Pro Leu His Trp Ala
35 40 45
Ala Gly Pro Glu Leu Gly Met Ala Pro His Leu Leu Trp Cys Pro
50 55 60
Thr Asn Gly Leu Gly Leu Gly Gly Ser Pro Ala Gly Gln Trp Gly
65 70 75
Gly Gly Ser His Tyr Arg Gly Leu Val Pro Gly Glu Pro Ala Gly
80 85 90
Arg Pro Pro Ala Leu Pro Leu Pro Gly Leu Ala Gly Leu Arg Asp
95 100 105
His Thr Gln Leu Leu Arg Met Ala Gly Gln Pro Trp Leu Ala Trp
110 115 120
Gly Thr Ala Ala Ala Arg Val Ser Ala Arg Pro Thr Arg Asp Cys
125 130 135
Ser Cys Thr Ser Arg Cys His His Ala Cys Asp Val Val Ala Val
140 145 150
Thr Leu Ser

<210> 7
<211> 369
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509244CD1

<400> 7

```
Met Ser Thr Lys Val Tyr Leu Asp Leu Glu Trp Thr Asp Tyr Arg
1               5                   10                  15
Leu Ser Trp Asp Pro Ala Glu His Asp Gly Ile Asp Ser Leu Arg
                20                  25                  30
Ile Thr Ala Glu Ser Val Trp Leu Pro Asp Val Val Leu Leu Asn
                35                  40                  45
Asn Asn Asp Gly Asn Phe Asp Val Ala Leu Asp Ile Ser Val Val
                50                  55                  60
Val Ser Ser Asp Gly Ser Val Arg Trp Gln Pro Pro Gly Ile Tyr
                65                  70                  75
Arg Ser Ser Cys Ser Ile Gln Val Thr Tyr Phe Pro Phe Asp Trp
                80                  85                  90
Gln Asn Cys Thr Met Val Phe Ser Ser Tyr Ser Tyr Asp Ser Ser
                95                  100                 105
Glu Val Ser Leu Gln Thr Gly Leu Gly Pro Asp Gly Gln Gly His
                110                 115                 120
Gln Glu Ile His Ile His Glu Gly Thr Phe Ile Glu Asn Gly Gln
                125                 130                 135
Trp Glu Ile Ile His Lys Pro Ser Arg Leu Ile Gln Pro Pro Gly
                140                 145                 150
Asp Pro Arg Gly Gly Arg Glu Gly Gln Arg Gln Glu Val Ile Phe
                155                 160                 165
Tyr Leu Ile Ile Arg Arg Lys Pro Leu Phe Tyr Leu Val Asn Val
                170                 175                 180
Ile Ala Pro Cys Ile Leu Ile Thr Leu Leu Ala Ile Phe Val Phe
                185                 190                 195
Tyr Leu Pro Pro Asp Ala Gly Glu Lys Met Gly Leu Ser Ile Phe
                200                 205                 210
Ala Leu Leu Thr Leu Thr Val Phe Leu Leu Leu Leu Ala Asp Lys
                215                 220                 225
Val Pro Glu Thr Ser Leu Ser Val Pro Ile Ile Ile Lys Tyr Leu
                230                 235                 240
Met Phe Thr Met Val Leu Val Thr Phe Ser Val Ile Leu Ser Val
                245                 250                 255
Val Val Leu Asn Leu His His Arg Ser Pro His Thr His Gln Met
                260                 265                 270
Pro Leu Trp Val Arg Gln Ile Phe Ile His Lys Leu Pro Leu Tyr
                275                 280                 285
Leu Arg Leu Lys Arg Pro Lys Pro Glu Arg Asp Leu Met Pro Glu
                290                 295                 300
Leu Arg Glu Val Val Ser Ser Ile Ser Tyr Ile Ala Arg Gln Leu
                305                 310                 315
Gln Glu Gln Glu Asp His Asp Ala Leu Lys Glu Asp Trp Gln Phe
                320                 325                 330
Val Ala Met Val Val Asp Arg Leu Phe Leu Trp Thr Phe Ile Ile
                335                 340                 345
Phe Thr Ser Val Gly Thr Leu Val Ile Phe Leu Asp Ala Thr Tyr
                350                 355                 360
His Leu Pro Pro Pro Asp Pro Phe Pro
                365
```

<210> 8
<211> 303
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509256CD1

<400> 8
Met Lys Phe Leu Leu Thr Thr Ala Phe Leu Ile Leu Ile Ser Leu
1 5 10 15
Trp Val Glu Glu Ala Tyr Ser Lys Glu Lys Ser Ser Lys Lys Gly
20 25 30
Lys Gly Lys Lys Lys Gln Tyr Leu Cys Pro Ser Gln Gln Ser Ala
35 40 45
Glu Asp Leu Ala Arg Val Pro Ala Asn Ser Thr Ser Asn Ile Leu
50 55 60
Asn Arg Leu Leu Val Ser Tyr Asp Pro Arg Ile Arg Pro Asn Phe
65 70 75
Lys Gly Ile Pro Val Asp Val Val Val Asn Ile Phe Ile Asn Ser
80 85 90
Phe Gly Ser Ile Gln Glu Thr Thr Met Asp Tyr Arg Val Asn Ile
95 100 105
Phe Leu Arg Gln Lys Trp Asn Asp Pro Arg Leu Lys Leu Pro Ser
110 115 120
Asp Phe Arg Gly Ser Asp Ala Leu Thr Val Asp Pro Thr Met Tyr
125 130 135
Lys Cys Leu Trp Lys Pro Asp Leu Phe Phe Ala Asn Glu Lys Ser
140 145 150
Ala Asn Phe His Asp Val Thr Gln Glu Asn Ile Leu Leu Phe Ile
155 160 165
Phe Arg Asp Gly Asp Val Leu Val Ser Met Arg Leu Ser Ile Thr
170 175 180
Leu Ser Cys Pro Leu Asp Leu Thr Leu Phe Pro Met Asp Thr Gln
185 190 195
Arg Cys Lys Met Gln Leu Glu Ser Phe Gly Tyr Thr Thr Asp Asp
200 205 210
Leu Arg Phe Ile Trp Gln Ser Gly Asp Pro Val Gln Leu Glu Lys
215 220 225
Ile Ala Leu Pro Gln Phe Asp Ile Lys Lys Glu Asp Ile Glu Tyr
230 235 240
Gly Asn Cys Thr Lys Tyr Tyr Lys Gly Thr Gly Tyr Tyr Thr Cys
245 250 255
Val Glu Val Ile Phe Thr Leu Arg Arg Gln Val Gly Phe Tyr Met
260 265 270
Met Gly Val Tyr Ala Pro Thr Leu Leu Ile Val Val Leu Ser Trp
275 280 285
Leu Ser Phe Trp Ile Asn Pro Asp Ala Ser Ala Ala Arg Val Pro
290 295 300
Leu Gly Trp

<210> 9
<211> 370
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509395CD1

<400> 9
Met Glu Pro Trp Pro Leu Leu Leu Leu Phe Ser Leu Cys Ser Ala
1 5 10 15
Gly Leu Val Leu Gly Ser Glu His Glu Thr Arg Leu Val Ala Lys
20 25 30
Leu Phe Lys Asp Tyr Ser Ser Val Val Arg Pro Val Glu Asp His
35 40 45
Arg Gln Val Val Glu Val Thr Val Gly Leu Gln Leu Ile Gln Leu
50 55 60
Ile Asn Val Asp Glu Val Asn Gln Ile Val Thr Thr Asn Asn Cys
65 70 75
Ser Met Lys Leu Gly Thr Trp Thr Tyr Asp Gly Ser Val Val Ala
80 85 90
Ile Asn Pro Glu Ser Asp Gln Pro Asp Leu Ser Asn Phe Met Glu
95 100 105
Ser Gly Glu Trp Val Ile Lys Glu Ser Arg Gly Trp Lys His Ser
110 115 120
Val Thr Tyr Ser Cys Cys Pro Asp Thr Pro Tyr Leu Asp Ile Thr
125 130 135
Tyr His Phe Val Met Gln Arg Leu Pro Leu Tyr Phe Ile Val Asn
140 145 150
Val Ile Ile Pro Cys Leu Leu Phe Ser Phe Leu Thr Gly Leu Val
155 160 165
Phe Tyr Leu Pro Thr Asp Ser Gly Glu Lys Met Thr Leu Ser Ile
170 175 180
Ser Val Leu Leu Ser Leu Thr Val Phe Leu Leu Val Ile Val Glu
185 190 195
Leu Ile Pro Ser Thr Ser Ser Ala Val Pro Leu Ile Gly Lys Tyr
200 205 210
Met Leu Phe Thr Met Val Phe Val Ile Ala Ser Ile Ile Ile Thr
215 220 225
Val Ile Val Ile Asn Thr His His Arg Ser Pro Ser Thr His Val
230 235 240
Met Pro Asn Trp Val Arg Lys Val Phe Ile Asp Thr Ile Pro Asn
245 250 255
Ile Met Phe Phe Ser Thr Met Lys Arg Pro Ser Arg Glu Lys Gln
260 265 270
Asp Lys Lys Ile Phe Thr Glu Asp Ile Asp Ile Ser Asp Ile Ser
275 280 285
Gly Lys Pro Gly Pro Pro Pro Met Gly Phe His Ser Pro Leu Ile
290 295 300
Lys His Pro Glu Val Lys Ser Ala Ile Glu Gly Ile Lys Tyr Ile
305 310 315
Ala Glu Thr Met Lys Ser Asp Gln Glu Ser Asn Asn Ala Ala Ala
320 325 330
Glu Trp Lys Tyr Val Ala Met Val Met Asp His Ile Leu Leu Gly
335 340 345
Val Phe Met Leu Val Cys Ile Ile Gly Thr Leu Ala Val Phe Ala
350 355 360
Gly Arg Leu Ile Glu Leu Asn Gln Gln Gly
365 370

<210> 10
<211> 283
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503287CD1

<400> 10
Met Glu Leu Lys Ala Glu Glu Glu Glu Val Gly Gly Val Gln Pro
1               5                   10                  15
Val Ser Ile Gln Ala Phe Ala Ser Ser Ser Thr Leu His Gly Leu
                20                  25                  30
Ala His Ile Phe Ser Tyr Glu Arg Leu Ser Leu Lys Arg Ala Leu
                35                  40                  45
Trp Ala Leu Cys Phe Leu Gly Ser Leu Ala Val Leu Leu Cys Val
                50                  55                  60
Cys Thr Glu Arg Val Gln Tyr Tyr Phe His Tyr His His Val Thr
                65                  70                  75
Lys Leu Asp Glu Val Ala Ala Ser Gln Leu Thr Phe Pro Ala Val
                80                  85                  90
Thr Leu Cys Asn Leu Asn Glu Phe Arg Phe Ser Gln Val Ser Lys
                95                  100                 105
Asn Asp Leu Tyr His Ala Gly Glu Leu Leu Ala Leu Leu Asn Asn
                110                 115                 120
Arg Tyr Glu Ile Pro Asp Thr Gln Met Ala Asp Glu Lys Gln Leu
                125                 130                 135
Glu Ile Leu Gln Asp Lys Ala Asn Phe Arg Ser Phe Lys Pro Lys
                140                 145                 150
Pro Phe Asn Met Arg Glu Phe Tyr Asp Arg Ala Gly His Asp Ile
                155                 160                 165
Arg Asp Met Leu Leu Ser Cys His Phe Arg Gly Glu Val Cys Ser
                170                 175                 180
Ala Glu Asp Phe Lys Val Val Phe Thr Arg Tyr Gly Lys Cys Tyr
                185                 190                 195
Thr Phe Asn Ser Gly Arg Asp Gly Arg Pro Arg Leu Lys Thr Met
                200                 205                 210
Lys Gly Gly Thr Gly Asn Gly Leu Glu Ile Met Leu Asp Ile Gln
                215                 220                 225
Gln Asp Glu Tyr Leu Pro Val Trp Gly Glu Thr Asp Glu Thr Ser
                230                 235                 240
Phe Glu Ala Gly Ile Lys Val Gln Ile Phe Pro Leu Val Cys Gly
                245                 250                 255
Lys Glu Gly Val Leu Thr Ile Glu Ser Ser Leu Cys Leu Tyr Pro
                260                 265                 270
Ile Leu Phe Thr Phe Asn Lys Thr Asn Leu Lys Lys Asn
                275                 280

<210> 11
<211> 90
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503320CD1

<400> 11
Met Arg Cys Ser Pro Gly Gly Val Trp Leu Ala Leu Ala Ala Ser
1               5                   10                  15

```
Leu Leu His Val Ser Leu Gln Gly Glu Phe Gln Arg Lys Leu Tyr
                20                  25                  30
Lys Glu Leu Val Lys Asn Tyr Asn Pro Leu Glu Arg Pro Val Ala
                35                  40                  45
Asn Asp Ser Gln Pro Leu Thr Val Tyr Phe Ser Leu Ser Leu Leu
                50                  55                  60
Gln Ile Met Asp Val Asp Glu Lys Asn Gln Val Leu Thr Thr Thr
                65                  70                  75
Thr Pro Thr Gly Ala Arg Cys Pro Ser Gly Pro Glu Ser Ser Phe
                80                  85                  90


<210> 12
<211> 549
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503335CD1

<400> 12
Met Pro Ala Cys Cys Ser Cys Ser Asp Val Phe Gln Tyr Glu Thr
  1               5                  10                  15
Asn Lys Val Thr Arg Ile Gln Ser Met Asn Tyr Gly Thr Ile Lys
                20                  25                  30
Trp Phe Phe His Val Ile Ile Phe Ser Tyr Val Cys Phe Ala Leu
                35                  40                  45
Val Ser Asp Lys Leu Tyr Gln Arg Lys Glu Pro Val Ile Ser Ser
                50                  55                  60
Val His Thr Lys Val Lys Gly Ile Ala Glu Val Lys Glu Glu Ile
                65                  70                  75
Val Glu Asn Gly Val Lys Lys Leu Val His Ser Val Phe Asp Thr
                80                  85                  90
Ala Asp Tyr Thr Phe Pro Leu Gln Gly Asn Ser Phe Phe Val Met
                95                 100                 105
Thr Asn Phe Leu Lys Thr Glu Gly Gln Glu Gln Arg Leu Cys Pro
               110                 115                 120
Glu Tyr Pro Thr Arg Arg Thr Leu Cys Ser Ser Asp Arg Gly Cys
               125                 130                 135
Lys Lys Gly Trp Met Asp Pro Gln Ser Lys Gly Ile Gln Thr Gly
               140                 145                 150
Arg Cys Val Val His Glu Gly Asn Gln Lys Thr Cys Glu Val Ser
               155                 160                 165
Ala Trp Cys Pro Ile Glu Ala Val Glu Glu Ala Pro Arg Pro Ala
               170                 175                 180
Leu Leu Asn Ser Ala Glu Asn Phe Thr Val Leu Ile Lys Asn Asn
               185                 190                 195
Ile Asp Phe Pro Gly His Asn Tyr Thr Thr Arg Asn Ile Leu Pro
               200                 205                 210
Gly Leu Asn Ile Thr Cys Thr Phe His Lys Thr Gln Asn Pro Gln
               215                 220                 225
Cys Pro Ile Phe Arg Leu Gly Asp Ile Phe Arg Glu Gln Ala Ile
               230                 235                 240
Ile Phe Gln Met Trp Gln Phe Arg Tyr Ala Lys Tyr Tyr Lys Glu
               245                 250                 255
```

Asn Asn Val Glu Lys Arg Thr Leu Ile Lys Val Phe Gly Ile Arg
260 265 270
Phe Asp Ile Leu Val Phe Gly Thr Gly Gly Lys Phe Asp Ile Ile
275 280 285
Gln Leu Val Val Tyr Ile Gly Ser Thr Leu Ser Tyr Phe Gly Leu
290 295 300
Ala Ala Val Phe Ile Asp Phe Leu Ile Asp Thr Tyr Ser Ser Asn
305 310 315
Cys Cys Arg Ser His Ile Tyr Pro Trp Cys Lys Cys Cys Gln Pro
320 325 330
Cys Val Val Asn Glu Tyr Tyr Tyr Arg Lys Lys Cys Glu Ser Ile
335 340 345
Val Glu Pro Lys Pro Thr Leu Lys Tyr Val Ser Phe Val Asp Glu
350 355 360
Ser His Ile Arg Met Val Asn Gln Gln Leu Leu Gly Arg Ser Leu
365 370 375
Gln Asp Val Lys Gly Gln Glu Val Pro Arg Pro Ala Met Asp Phe
380 385 390
Thr Asp Leu Ser Arg Leu Pro Leu Ala Leu His Asp Thr Pro Pro
395 400 405
Ile Pro Gly Gln Pro Glu Glu Ile Gln Leu Leu Arg Lys Glu Ala
410 415 420
Thr Pro Arg Ser Arg Asp Ser Pro Val Trp Cys Gln Cys Gly Ser
425 430 435
Cys Leu Pro Ser Gln Leu Pro Glu Ser His Arg Cys Leu Glu Glu
440 445 450
Leu Cys Cys Arg Lys Lys Pro Gly Ala Cys Ile Thr Thr Ser Glu
455 460 465
Leu Phe Arg Lys Leu Val Leu Ser Arg His Val Leu Gln Phe Leu
470 475 480
Leu Leu Tyr Gln Glu Pro Leu Leu Ala Leu Asp Val Asp Ser Thr
485 490 495
Asn Ser Arg Leu Arg His Cys Ala Tyr Arg Cys Tyr Ala Thr Trp
500 505 510
Arg Phe Gly Ser Gln Asp Met Ala Asp Phe Ala Ile Leu Pro Ser
515 520 525
Cys Cys Arg Trp Arg Ile Arg Lys Glu Phe Pro Lys Ser Glu Gly
530 535 540
Gln Tyr Ser Gly Phe Lys Ser Pro Tyr
545

<210> 13
<211> 246
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503952CD1

<400> 13
Met Leu Ser Ser Val Met Ala Pro Leu Trp Ala Cys Ile Leu Val
1 5 10 15
Ala Ala Gly Ile Leu Ala Thr Asp Thr His His Pro Gln Asp Ser
20 25 30
Ala Leu Tyr His Leu Ser Lys Gln Leu Leu Gln Lys Tyr His Lys

```
                    35                  40                  45
Glu Val Arg Pro Val Tyr Asn Trp Thr Lys Ala Thr Thr Val Tyr
                    50                  55                  60
Leu Asp Leu Phe Val His Ala Ile Leu Asp Val Asp Ala Glu Asn
                    65                  70                  75
Gln Ile Leu Lys Thr Ser Val Trp Tyr Gln Glu Val Trp Asn Asp
                    80                  85                  90
Glu Phe Leu Ser Trp Asn Ser Ser Met Phe Asp Glu Ile Arg Glu
                    95                  100                 105
Ile Ser Leu Pro Leu Ser Ala Ile Trp Ala Pro Asp Ile Ile Ile
                    110                 115                 120
Asn Glu Phe Val Asp Ile Glu Arg Tyr Pro Asp Leu Pro Tyr Val
                    125                 130                 135
Tyr Val Asn Ser Ser Gly Thr Ile Glu Asn Tyr Lys Pro Ile Gln
                    140                 145                 150
Val Val Ser Ala Cys Ser Leu Glu Thr Tyr Ala Phe Pro Phe Asp
                    155                 160                 165
Val Gln Asn Cys Ser Leu Thr Phe Lys Ser Ile Leu His Thr Val
                    170                 175                 180
Glu Asp Val Asp Leu Ala Phe Leu Arg Ser Pro Glu Asp Ile Gln
                    185                 190                 195
His Asp Lys Lys Ala Phe Leu Asn Asp Ser Glu Trp Glu Leu Leu
                    200                 205                 210
Ser Val Ser Ser Thr Tyr Ser Ile Leu Gln Ser Ser Ala Gly Gly
                    215                 220                 225
Phe Ala Gln Ile Gln Phe Asn Gly Thr Ser Ser Pro Ser Ala Trp
                    230                 235                 240
Pro Ser Trp Phe Ser Ala
                    245

<210> 14
<211> 273
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504530CD1

<400> 14
Met Val Gln Ala Ser Gly His Arg Arg Ser Thr Arg Gly Ser Lys
  1               5                  10                  15
Met Val Ser Trp Ser Val Ile Ala Lys Ile Gln Glu Ile Leu Gln
                    20                  25                  30
Arg Lys Met Val Arg Glu Phe Leu Ala Glu Phe Met Ser Thr Tyr
                    35                  40                  45
Val Met Met Val Phe Gly Leu Gly Ser Val Ala His Met Val Leu
                    50                  55                  60
Asn Lys Lys Tyr Gly Ser Tyr Leu Gly Val Asn Leu Gly Phe Gly
                    65                  70                  75
Phe Gly Val Thr Met Gly Val His Val Ala Gly Arg Ile Ser Gly
                    80                  85                  90
Ala His Met Asn Ala Ala Val Thr Phe Ala Asn Cys Ala Leu Gly
                    95                  100                 105
Arg Val Pro Trp Arg Lys Phe Pro Val Tyr Val Leu Gly Gln Phe
                    110                 115                 120
```

Leu Gly Ser Phe Leu Ala Ala Ala Thr Ile Tyr Ser Leu Phe Tyr
125 130 135
Thr Ala Ile Leu His Phe Ser Gly Gly Gln Leu Met Val Thr Gly
140 145 150
Pro Val Ala Thr Ala Gly Ile Phe Ala Thr Tyr Leu Pro Asp His
155 160 165
Met Thr Leu Trp Arg Gly Phe Leu Asn Glu Ala Trp Leu Thr Gly
170 175 180
Met Leu Gln Leu Cys Leu Phe Ala Ile Thr Asp Gln Glu Asn Asn
185 190 195
Pro Ala Leu Pro Gly Thr Glu Ala Leu Val Ile Gly Ile Leu Val
200 205 210
Val Ile Ile Gly Val Ser Leu Gly Met Asn Thr Gly Tyr Ala Ile
215 220 225
Asn Pro Ser Arg Asp Leu Pro Pro Arg Ile Phe Thr Phe Ile Ala
230 235 240
Gly Trp Gly Lys Gln Val Phe Arg Trp His His Leu Pro Gly Leu
245 250 255
His Trp Leu His His Pro Thr Gly Ala Pro Glu Ile Gly Gly Phe
260 265 270
Cys Gly Val

<210> 15
<211> 245
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509303CD1

<400> 15
Met Glu Gly Asn Lys Leu Glu Glu Gln Asp Ser Ser Pro Pro Gln
1 5 10 15
Ser Thr Pro Gly Leu Met Lys Gly Asn Lys Arg Glu Glu Gln Gly
20 25 30
Leu Gly Pro Glu Pro Ala Ala Pro Gln Gln Pro Thr Ala Glu Glu
35 40 45
Glu Ala Leu Ile Glu Phe His Arg Ser Tyr Arg Glu Leu Phe Glu
50 55 60
Phe Phe Cys Asn Asn Thr Thr Ile His Gly Ala Ile Arg Leu Val
65 70 75
Cys Ser Gln His Asn Arg Met Lys Thr Ala Phe Trp Ala Val Leu
80 85 90
Trp Leu Cys Thr Phe Gly Met Met Tyr Trp Gln Phe Gly Leu Leu
95 100 105
Phe Gly Glu Tyr Phe Ser Tyr Pro Val Ser Leu Asn Ile Asn Leu
110 115 120
Asn Ser Asp Lys Leu Val Phe Pro Ala Val Thr Ile Cys Thr Leu
125 130 135
Asn Pro Tyr Arg Tyr Pro Glu Ile Lys Glu Glu Leu Glu Glu Leu
140 145 150
Asp Arg Ile Thr Glu Gln Thr Leu Phe Asp Leu Tyr Lys Tyr Ser
155 160 165
Ser Phe Thr Thr Leu Val Ala Gly Ser Arg Ser Arg Arg Asp Leu

```
                170                 175                 180
Arg Gly Thr Leu Pro His Pro Leu Gln Arg Leu Arg Val Pro Pro
                185                 190                 195
Pro Pro His Gly Ala Arg Arg Ala Arg Ser Val Ala Ser Ser Leu
                200                 205                 210
Arg Asp Asn Asn Pro Gln Val Asp Trp Lys Asp Trp Lys Ile Gly
                215                 220                 225
Phe Gln Leu Glu Leu Leu Ser Leu Pro Pro Pro Asp Val Trp Lys
                230                 235                 240
Leu Leu Tyr Phe Gln
                245

<210> 16
<211> 364
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509910CD1

<400> 16
Met Pro Ala Cys Cys Ser Cys Ser Asp Val Phe Gln Tyr Glu Thr
  1               5                  10                  15
Asn Lys Val Thr Arg Ile Gln Ser Met Asn Tyr Gly Thr Ile Lys
                 20                  25                  30
Trp Phe Phe His Val Ile Ile Phe Ser Tyr Val Cys Phe Ala Leu
                 35                  40                  45
Val Ser Asp Lys Leu Tyr Gln Arg Lys Glu Pro Val Ile Ser Ser
                 50                  55                  60
Val His Thr Lys Val Lys Gly Ile Ala Glu Val Lys Glu Glu Ile
                 65                  70                  75
Val Glu Asn Gly Val Lys Lys Leu Val His Ser Val Phe Asp Thr
                 80                  85                  90
Ala Asp Tyr Thr Phe Pro Leu Gln Gly Asn Ser Phe Phe Val Met
                 95                 100                 105
Thr Asn Phe Leu Lys Thr Glu Gly Gln Glu Gln Arg Leu Cys Pro
                110                 115                 120
Glu Tyr Pro Thr Arg Arg Thr Leu Cys Ser Ser Asp Arg Gly Cys
                125                 130                 135
Lys Lys Gly Trp Met Asp Pro Gln Ser Lys Gly Ile Gln Thr Gly
                140                 145                 150
Arg Cys Val Val His Glu Gly Asn Gln Lys Thr Cys Glu Val Ser
                155                 160                 165
Ala Trp Cys Pro Ile Glu Ala Val Glu Glu Ala Pro Arg Pro Ala
                170                 175                 180
Leu Leu Asn Ser Ala Glu Asn Phe Thr Val Leu Ile Lys Asn Asn
                185                 190                 195
Ile Asp Phe Pro Gly His Asn Tyr Thr Thr Arg Asn Ile Leu Pro
                200                 205                 210
Gly Leu Asn Ile Thr Cys Thr Phe His Lys Thr Gln Asn Pro Gln
                215                 220                 225
Cys Pro Ile Phe Arg Leu Gly Asp Ile Phe Arg Glu Thr Gly Asp
                230                 235                 240
Asn Phe Ser Asp Val Ala Ile Gln Gly Gly Ile Met Gly Ile Glu
                245                 250                 255
```

```
Ile Tyr Trp Asp Cys Asn Leu Asp Arg Trp Phe His His Cys Arg
                260                 265                 270
Pro Lys Tyr Ser Phe Arg Arg Leu Asp Asp Lys Thr Thr Asn Val
                275                 280                 285
Ser Leu Tyr Pro Gly Tyr Asn Phe Arg Tyr Ala Lys Tyr Tyr Lys
                290                 295                 300
Glu Asn Asn Val Glu Lys Arg Thr Leu Ile Lys Val Phe Gly Ile
                305                 310                 315
Arg Phe Asp Ile Leu Val Phe Gly Thr Gly Gly Lys Phe Asp Ile
                320                 325                 330
Ile Gln Leu Val Val Tyr Ile Gly Ser Thr Leu Ser Tyr Phe Gly
                335                 340                 345
Leu Val Arg Asp Ser Leu Phe His Ala Leu Gly Lys Trp Phe Gly
                350                 355                 360
Glu Gly Ser Asp


<210> 17
<211> 1623
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509982CD1

<400> 17
Met Asn Met Lys Gln Lys Ser Val Tyr Gln Gln Thr Lys Ala Leu
1               5                   10                  15
Leu Cys Lys Asn Phe Leu Lys Lys Trp Arg Met Lys Arg Glu Ser
                20                  25                  30
Leu Leu Glu Trp Gly Leu Ser Ile Leu Leu Gly Leu Cys Ile Ala
                35                  40                  45
Leu Phe Ser Ser Ser Met Arg Asn Val Gln Phe Pro Gly Met Ala
                50                  55                  60
Pro Gln Asn Leu Gly Arg Val Asp Lys Phe Asn Ser Ser Ser Leu
                65                  70                  75
Met Val Val Tyr Thr Pro Ile Ser Asn Leu Thr Gln Gln Ile Met
                80                  85                  90
Asn Lys Thr Ala Leu Ala Pro Leu Leu Lys Gly Thr Ser Val Ile
                95                  100                 105
Gly Ala Pro Asn Lys Thr His Met Asp Glu Ile Leu Leu Glu Asn
                110                 115                 120
Leu Pro Tyr Ala Met Gly Ile Ile Phe Asn Glu Thr Phe Ser Tyr
                125                 130                 135
Lys Leu Ile Phe Phe Gln Gly Tyr Asn Ser Pro Leu Trp Lys Glu
                140                 145                 150
Asp Phe Ser Ala His Cys Trp Asp Gly Tyr Gly Glu Phe Ser Cys
                155                 160                 165
Thr Leu Thr Lys Tyr Trp Asn Arg Gly Phe Val Ala Leu Gln Thr
                170                 175                 180
Ala Ile Asn Thr Ala Ile Ile Glu Ile Thr Thr Asn His Pro Val
                185                 190                 195
Met Glu Glu Leu Met Ser Val Thr Ala Ile Thr Met Lys Thr Leu
                200                 205                 210
Pro Phe Ile Thr Lys Asn Leu Leu His Asn Glu Met Phe Ile Leu
```

215 220 225
Phe Phe Leu Leu His Phe Ser Pro Leu Val Tyr Phe Ile Ser Leu
230 235 240
Asn Val Thr Lys Glu Arg Lys Lys Ser Lys Asn Leu Met Lys Met
245 250 255
Met Gly Leu Gln Asp Ser Ala Phe Trp Leu Ser Trp Gly Leu Ile
260 265 270
Tyr Ala Gly Phe Ile Phe Ile Ile Ser Ile Phe Ile Thr Ile Ile
275 280 285
Ile Thr Phe Thr Gln Ile Ile Val Met Thr Gly Phe Met Val Ile
290 295 300
Phe Ile Pro Phe Phe Leu Tyr Gly Leu Ser Leu Val Ala Leu Val
305 310 315
Phe Leu Leu Ser Val Leu Leu Lys Lys Ala Val Leu Thr Asn Leu
320 325 330
Val Val Phe Leu Leu Thr Leu Phe Trp Gly Cys Leu Gly Phe Thr
335 340 345
Val Phe Tyr Glu Gln Leu Pro Ser Ser Leu Glu Trp Ile Leu Asn
350 355 360
Ile Cys Ser Pro Phe Ala Phe Thr Thr Gly Met Ile Gln Ile Ile
365 370 375
Lys Leu Asp Tyr Asn Leu Asn Gly Val Ile Phe Pro Asp Pro Ser
380 385 390
Gly Asp Ser Tyr Thr Met Ile Ala Thr Phe Ser Met Leu Leu Leu
395 400 405
Asp Gly Leu Ile Tyr Leu Leu Leu Ala Leu Tyr Phe Asp Lys Ile
410 415 420
Leu Pro Tyr Gly Asp Glu Arg His Tyr Ser Pro Leu Phe Phe Leu
425 430 435
Asn Ser Ser Ser Cys Phe Gln His Gln Arg Thr Asn Ala Lys Val
440 445 450
Ile Glu Lys Glu Ile Asp Ala Glu His Pro Ser Asp Asp Tyr Phe
455 460 465
Glu Pro Val Ala Pro Glu Phe Gln Gly Lys Glu Ala Ile Arg Ile
470 475 480
Arg Asn Val Lys Lys Glu Tyr Lys Gly Lys Ser Gly Lys Val Glu
485 490 495
Ala Leu Lys Gly Leu Leu Phe Asp Ile Tyr Glu Gly Gln Ile Thr
500 505 510
Ala Ile Leu Gly His Ser Gly Ala Gly Lys Ser Ser Leu Leu Asn
515 520 525
Ile Leu Asn Gly Leu Ser Val Pro Thr Glu Gly Ser Val Thr Ile
530 535 540
Tyr Asn Lys Asn Leu Ser Glu Met Gln Asp Leu Glu Glu Ile Arg
545 550 555
Lys Ile Thr Gly Val Cys Pro Gln Phe Asn Val Gln Phe Asp Ile
560 565 570
Leu Thr Val Lys Glu Asn Leu Ser Leu Phe Ala Lys Ile Lys Gly
575 580 585
Ile His Leu Lys Glu Val Glu Gln Glu Val Gln Arg Ile Leu Leu
590 595 600
Glu Leu Asp Met Gln Asn Ile Gln Asp Asn Leu Ala Lys His Leu
605 610 615
Ser Glu Gly Gln Lys Arg Lys Leu Thr Phe Gly Ile Thr Ile Leu
620 625 630
Gly Asp Pro Gln Ile Leu Leu Leu Asp Glu Pro Thr Thr Gly Leu

635 640 645
Asp Pro Phe Ser Arg Asp Gln Val Trp Ser Leu Leu Arg Glu Arg
650 655 660
Arg Ala Asp His Val Ile Leu Phe Ser Thr Gln Ser Met Asp Glu
665 670 675
Ala Asp Ile Leu Ala Asp Arg Lys Val Ile Met Ser Asn Gly Arg
680 685 690
Leu Lys Cys Ala Gly Ser Ser Met Phe Leu Lys Arg Arg Trp Gly
695 700 705
Leu Gly Tyr His Leu Ser Leu His Arg Asn Glu Ile Cys Asn Pro
710 715 720
Glu Gln Ile Thr Ser Phe Ile Thr His His Ile Pro Asp Ala Lys
725 730 735
Leu Lys Thr Glu Asn Lys Glu Lys Leu Val Tyr Thr Leu Pro Leu
740 745 750
Glu Arg Thr Asn Thr Phe Pro Asp Leu Phe Ser Asp Leu Asp Lys
755 760 765
Cys Ser Asp Gln Gly Val Thr Gly Tyr Asp Ile Ser Met Ser Thr
770 775 780
Leu Asn Glu Val Phe Met Lys Leu Glu Gly Gln Ser Thr Ile Glu
785 790 795
Gln Gly Lys Ala Ile Cys Ile Asn Phe Glu Gln Val Glu Met Ile
800 805 810
Arg Asp Ser Glu Ser Leu Asn Glu Met Glu Leu Ala His Ser Ser
815 820 825
Phe Ser Glu Met Gln Thr Ala Val Ser Asp Met Gly Leu Trp Arg
830 835 840
Met Gln Val Phe Ala Met Ala Arg Leu Arg Phe Leu Lys Leu Lys
845 850 855
Arg Gln Thr Lys Val Leu Leu Thr Leu Leu Leu Val Phe Gly Ile
860 865 870
Ala Ile Phe Pro Leu Ile Val Glu Asn Ile Ile Tyr Ala Met Leu
875 880 885
Asn Glu Lys Ile Asp Trp Glu Phe Lys Asn Glu Leu Tyr Phe Leu
890 895 900
Ser Pro Gly Gln Leu Pro Gln Glu Pro Arg Thr Ser Leu Leu Ile
905 910 915
Ile Asn Asn Thr Glu Ser Asn Ile Glu Asp Phe Ile Lys Ser Leu
920 925 930
Lys His Gln Asn Ile Leu Leu Glu Val Asp Asp Phe Glu Asn Arg
935 940 945
Asn Gly Thr Asp Gly Leu Ser Tyr Asn Gly Ala Ile Ile Val Ser
950 955 960
Gly Lys Gln Lys Asp Tyr Arg Phe Ser Val Val Cys Asn Thr Lys
965 970 975
Arg Leu His Cys Phe Pro Ile Leu Met Asn Ile Ile Ser Asn Gly
980 985 990
Leu Leu Gln Met Phe Asn His Thr Gln His Ile Arg Ile Glu Ser
995 1000 1005
Ser Pro Phe Pro Leu Ser His Ile Gly Leu Trp Thr Gly Leu Pro
1010 1015 1020
Asp Gly Ser Phe Phe Leu Phe Leu Val Leu Cys Ser Ile Ser Pro
1025 1030 1035
Tyr Ile Thr Met Gly Ser Ile Ser Asp Tyr Lys Lys Asn Ala Lys
1040 1045 1050
Ser Gln Leu Trp Ile Ser Gly Leu Tyr Thr Ser Ala Tyr Trp Cys

1055 1060 1065
Gly Gln Ala Leu Val Asp Val Ser Phe Phe Ile Leu Ile Leu Leu
1070 1075 1080
Leu Met Tyr Leu Ile Phe Tyr Ile Glu Asn Met Gln Tyr Leu Leu
1085 1090 1095
Ile Thr Ser Gln Ile Val Phe Ala Leu Val Ile Val Thr Pro Gly
1100 1105 1110
Tyr Ala Ala Ser Leu Val Phe Phe Ile Tyr Met Ile Ser Phe Ile
1115 1120 1125
Phe Arg Lys Arg Arg Lys Asn Ser Gly Leu Trp Ser Phe Tyr Phe
1130 1135 1140
Phe Phe Ala Ser Thr Ile Met Phe Ser Ile Thr Leu Ile Asn His
1145 1150 1155
Phe Asp Leu Ser Ile Leu Ile Thr Thr Met Val Leu Val Pro Ser
1160 1165 1170
Tyr Thr Leu Leu Gly Phe Lys Thr Phe Leu Glu Val Arg Asp Gln
1175 1180 1185
Glu His Tyr Arg Glu Phe Pro Glu Ala Asn Phe Glu Leu Ser Ala
1190 1195 1200
Thr Asp Phe Leu Val Cys Phe Ile Pro Tyr Phe Gln Thr Leu Leu
1205 1210 1215
Phe Val Phe Val Leu Arg Tyr Met Glu Leu Lys Cys Gly Lys Lys
1220 1225 1230
Arg Met Arg Lys Asp Pro Val Phe Arg Ile Ser Pro Gln Ser Arg
1235 1240 1245
Asp Ala Lys Pro Asn Pro Glu Glu Pro Ile Asp Glu Asp Glu Asp
1250 1255 1260
Ile Gln Thr Glu Arg Ile Arg Thr Val Thr Ala Leu Thr Thr Ser
1265 1270 1275
Ile Leu Asp Glu Lys Pro Val Ile Ile Ala Ser Cys Leu His Lys
1280 1285 1290
Glu Tyr Ala Gly Gln Lys Lys Ser Cys Phe Ser Lys Arg Lys Lys
1295 1300 1305
Lys Ile Ala Ala Arg Asn Ile Ser Phe Cys Val Gln Glu Gly Glu
1310 1315 1320
Ile Leu Gly Leu Leu Gly Pro Ser Gly Ala Gly Lys Ser Ser Ser
1325 1330 1335
Ile Arg Met Ile Ser Gly Ile Thr Lys Pro Thr Ala Gly Glu Val
1340 1345 1350
Glu Leu Lys Gly Cys Ser Ser Val Leu Gly His Leu Gly Tyr Cys
1355 1360 1365
Pro Gln Glu Asn Val Leu Trp Pro Met Leu Thr Leu Arg Glu His
1370 1375 1380
Leu Glu Val Tyr Ala Ala Val Lys Gly Leu Arg Glu Ala Asp Ala
1385 1390 1395
Arg Leu Ala Ile Ala Arg Leu Val Ser Ala Phe Lys Leu His Glu
1400 1405 1410
Gln Leu Asn Val Pro Val Gln Lys Leu Thr Ala Gly Ile Thr Arg
1415 1420 1425
Lys Leu Cys Phe Val Leu Ser Leu Leu Gly Asn Ser Pro Val Leu
1430 1435 1440
Leu Leu Asp Glu Pro Ser Thr Gly Ile Asp Pro Thr Gly Gln Gln
1445 1450 1455
Gln Met Trp Gln Ala Ile Gln Ala Val Val Lys Asn Thr Glu Arg
1460 1465 1470
Gly Val Leu Leu Thr Thr His Asn Leu Ala Glu Ala Glu Ala Leu

1475 1480 1485
Cys Asp Arg Val Ala Ile Met Val Ser Gly Arg Leu Arg Cys Ile
1490 1495 1500
Gly Ser Ile Gln His Leu Lys Asn Lys Leu Gly Lys Asp Tyr Ile
1505 1510 1515
Leu Glu Leu Lys Val Lys Glu Thr Ser Gln Val Thr Leu Val His
1520 1525 1530
Thr Glu Ile Leu Lys Leu Phe Pro Gln Ala Ala Gly Gln Glu Arg
1535 1540 1545
Tyr Ser Ser Leu Leu Thr Tyr Lys Leu Pro Val Ala Asp Val Tyr
1550 1555 1560
Pro Leu Ser Gln Thr Phe His Lys Leu Glu Ala Val Lys His Asn
1565 1570 1575
Phe Asn Leu Glu Glu Tyr Ser Leu Ser Gln Cys Thr Leu Glu Lys
1580 1585 1590
Val Phe Leu Glu Leu Ser Lys Glu Gln Glu Val Gly Asn Phe Asp
1595 1600 1605
Glu Glu Ile Asp Thr Thr Met Arg Trp Lys Leu Leu Pro His Ser
1610 1615 1620
Asp Glu Pro

<210> 18
<211> 611
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510082CD1

<400> 18
Met Pro Ala Pro Arg Ala Arg Glu Gln Pro Arg Val Pro Gly Glu
1 5 10 15
Arg Gln Pro Leu Leu Pro Arg Gly Ala Arg Gly Pro Arg Arg Trp
20 25 30
Arg Arg Ala Ala Gly Ala Ala Val Leu Leu Val Glu Met Leu Glu
35 40 45
Arg Ala Ala Phe Phe Gly Val Thr Ala Asn Leu Val Leu Tyr Leu
50 55 60
Asn Ser Thr Asn Phe Asn Trp Thr Gly Glu Gln Ala Thr Arg Ala
65 70 75
Ala Leu Val Phe Leu Gly Ala Ser Tyr Leu Leu Ala Pro Val Gly
80 85 90
Gly Trp Leu Ala Asp Val Tyr Leu Gly Arg Tyr Arg Ala Val Ala
95 100 105
Leu Ser Leu Leu Leu Tyr Leu Ala Ala Ser Gly Leu Leu Pro Ala
110 115 120
Thr Ala Phe Pro Asp Gly Arg Ser Ser Phe Cys Gly Glu Met Pro
125 130 135
Ala Ser Pro Leu Gly Pro Ala Cys Pro Ser Ala Gly Cys Pro Arg
140 145 150
Ser Ser Pro Ser Pro Tyr Cys Ala Pro Val Leu Tyr Ala Gly Leu
155 160 165
Leu Leu Leu Gly Leu Ala Ala Ser Ser Val Arg Ser Asn Leu Thr
170 175 180

Ser Phe Gly Ala Asp Gln Val Met Asp Leu Gly Arg Asp Ala Thr
185 190 195
Arg Arg Phe Phe Asn Trp Phe Tyr Trp Ser Ile Asn Leu Gly Ala
200 205 210
Val Leu Ser Leu Leu Val Val Ala Phe Ile Gln Gln Asn Ile Ser
215 220 225
Phe Leu Leu Gly Tyr Ser Ile Pro Val Gly Cys Val Gly Leu Ala
230 235 240
Phe Phe Ile Phe Leu Phe Ala Thr Pro Val Phe Ile Thr Lys Pro
245 250 255
Pro Met Gly Ser Gln Val Ser Ser Met Leu Lys Leu Ala Leu Gln
260 265 270
Asn Cys Cys Pro Gln Leu Trp Gln Arg His Ser Ala Arg Asp Arg
275 280 285
Gln Cys Ala Arg Val Leu Ala Asp Glu Arg Ser Pro Gln Pro Gly
290 295 300
Ala Ser Pro Gln Glu Asp Ile Ala Asn Phe Gln Val Leu Val Lys
305 310 315
Ile Leu Pro Val Met Val Thr Leu Val Pro Tyr Trp Met Val Tyr
320 325 330
Phe Gln Met Gln Ser Thr Tyr Val Leu Gln Gly Leu His Leu His
335 340 345
Ile Pro Asn Ile Phe Pro Ala Asn Pro Ala Asn Ile Ser Val Ala
350 355 360
Leu Arg Ala Gln Gly Ser Ser Tyr Thr Ile Pro Glu Ala Trp Leu
365 370 375
Leu Leu Ala Asn Val Val Val Val Leu Ile Leu Val Pro Leu Lys
380 385 390
Asp Arg Leu Ile Asp Pro Leu Leu Leu Arg Cys Lys Leu Leu Pro
395 400 405
Ser Ala Leu Gln Lys Met Ala Leu Gly Met Phe Phe Gly Phe Thr
410 415 420
Ser Val Ile Val Ala Gly Val Leu Glu Met Glu Arg Leu His Tyr
425 430 435
Ile His His Asn Glu Thr Val Ser Gln Gln Ile Gly Glu Val Leu
440 445 450
Tyr Asn Ala Ala Pro Leu Ser Ile Trp Trp Gln Ile Pro Gln Tyr
455 460 465
Leu Leu Ile Gly Ile Ser Glu Ile Phe Ala Ser Ile Pro Gly Leu
470 475 480
Glu Phe Ala Tyr Ser Glu Ala Pro Arg Ser Met Gln Gly Ala Ile
485 490 495
Met Gly Ile Phe Phe Cys Leu Ser Gly Val Gly Ser Leu Leu Gly
500 505 510
Ser Ser Leu Val Ala Leu Leu Ser Leu Pro Gly Gly Trp Leu His
515 520 525
Cys Pro Lys Asp Phe Gly Asn Ile Asn Asn Cys Arg Met Asp Leu
530 535 540
Tyr Phe Phe Leu Leu Ala Gly Ile Gln Ala Val Thr Ala Leu Leu
545 550 555
Phe Val Trp Ile Ala Gly Arg Tyr Glu Arg Ala Ser Gln Gly Pro
560 565 570
Ala Ser His Arg Pro Phe Gln His Gly Gln Gly Leu Asp Arg Pro
575 580 585
Tyr Pro Gly Pro Leu Val Tyr Ser Thr Gly Lys Asn Gly Ser Ser
590 595 600

Pro Ser Ser Gly Phe Leu Leu Gly Leu Phe Cys
605 610

<210> 19
<211> 55
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510367CD1

<400> 19
Met Thr Gly Gln Gly Gln Ser Ala Ser Gly Ser Ser Ala Trp Ser
1 5 10 15
Thr Val Phe Arg His Val Arg Tyr Glu Asn Leu Ile Ala Gly Val
20 25 30
Ser Gly Gly Val Leu Ser Asn Leu Ala Leu His Pro Leu Asp Leu
35 40 45
Val Lys Ile Arg Phe Ala Gly Thr Ile Leu
50 55

<210> 20
<211> 287
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510413CD1

<400> 20
Met Asp Met Ala Trp Gln Met Met Gln Leu Leu Leu Leu Ala Leu
1 5 10 15
Val Thr Ala Ala Gly Ser Ala Gln Pro Arg Ser Ala Arg Ala Arg
20 25 30
Thr Asp Leu Leu Asn Val Cys Met Asn Ala Lys His His Lys Thr
35 40 45
Gln Pro Ser Pro Glu Asp Glu Leu Tyr Gly Gln Val Gly Ala Pro
50 55 60
Gln Gly Pro Ser Pro Gly Ser Val Pro Leu Asp Asp Leu Pro Gly
65 70 75
Ala Glu Glu Pro Glu Tyr Gly Gly Asp Gly Cys Gly Gly Glu Arg
80 85 90
Leu Ser Pro Val Ser Ser Pro Pro Ser Ala Val Pro Gly Arg Arg
95 100 105
Met Pro Ala Ala Arg Pro Ala Pro Ala Arg Ser Cys Thr Arg Thr
110 115 120
Pro Pro Ala Cys Thr Thr Leu Thr Gly Ile Thr Val Val Arg Trp
125 130 135
Asn Pro Pro Ala Ser Ala Thr Leu Ser Arg Thr Ala Val Ser Glu
140 145 150
Cys Ser Pro Asn Leu Gly Pro Trp Ile Arg Gln Val Asn Gln Ser
155 160 165
Trp Arg Lys Glu Arg Ile Leu Asn Val Pro Leu Cys Lys Glu Asp
170 175 180

```
Cys Glu Arg Trp Trp Glu Asp Cys Arg Thr Ser Tyr Thr Cys Lys
                185                 190                 195
Ser Asn Trp His Lys Gly Trp Asn Trp Thr Ser Gly Ile Asn Glu
                200                 205                 210
Cys Pro Ala Gly Ala Leu Cys Ser Thr Phe Glu Ser Tyr Phe Pro
                215                 220                 225
Thr Pro Ala Ala Leu Cys Glu Gly Leu Trp Ser His Ser Phe Lys
                230                 235                 240
Val Ser Asn Tyr Ser Arg Gly Ser Gly Arg Cys Ile Gln Met Trp
                245                 250                 255
Phe Asp Ser Ala Gln Gly Asn Pro Asn Glu Glu Val Ala Lys Phe
                260                 265                 270
Tyr Ala Ala Ala Met Asn Ala Gly Ala Pro Ser Arg Gly Ile Ile
                275                 280                 285
Asp Ser


<210> 21
<211> 55
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1721303CD1

<400> 21
Met Ala Ser Val Gly Glu Cys Pro Ala Pro Val Pro Val Lys Asp
  1               5                  10                  15
Lys Lys Leu Leu Glu Val Lys Leu Gly Glu Leu Pro Ser Trp Ile
                 20                  25                  30
Leu Met Arg Asp Phe Ser Pro Ser Gly Ile Phe Gly Ala Phe Gln
                 35                  40                  45
Arg Glu His Glu Arg Leu Arg Lys Tyr His
                 50                  55

<210> 22
<211> 272
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7502007CD1

<400> 22
Met Gly Ser Gly His Cys Leu Arg Ser Thr Arg Gly Ser Lys Met
  1               5                  10                  15
Val Ser Trp Ser Val Ile Ala Lys Ile Gln Glu Ile Leu Gln Arg
                 20                  25                  30
Lys Met Val Arg Glu Phe Leu Ala Glu Phe Met Ser Thr Tyr Val
                 35                  40                  45
Met Met Val Phe Gly Leu Gly Ser Val Ala His Met Val Leu Asn
                 50                  55                  60
Lys Lys Tyr Gly Ser Tyr Leu Gly Val Asn Leu Gly Phe Gly Phe
                 65                  70                  75
```

Gly Val Thr Met Gly Val His Val Ala Gly Arg Ile Ser Gly Ala
80 85 90
His Met Asn Ala Ala Val Thr Phe Ala Asn Cys Ala Leu Gly Arg
95 100 105
Val Pro Trp Arg Lys Phe Pro Val Tyr Val Leu Gly Gln Phe Leu
110 115 120
Gly Ser Phe Leu Ala Ala Ala Thr Ile Tyr Ser Leu Phe Tyr Thr
125 130 135
Ala Ile Leu His Phe Ser Gly Gly Gln Leu Met Val Thr Gly Pro
140 145 150
Val Ala Thr Ala Gly Ile Phe Ala Thr Tyr Leu Pro Asp His Met
155 160 165
Thr Leu Trp Arg Gly Phe Leu Asn Glu Ala Trp Leu Thr Gly Met
170 175 180
Leu Gln Leu Cys Leu Phe Ala Ile Thr Asp Gln Glu Asn Asn Pro
185 190 195
Ala Leu Pro Gly Thr Glu Ala Leu Val Ile Gly Ile Leu Val Val
200 205 210
Ile Ile Gly Val Ser Leu Gly Met Asn Thr Gly Tyr Ala Ile Asn
215 220 225
Pro Ser Arg Asp Leu Pro Pro Arg Ile Phe Thr Phe Ile Ala Gly
230 235 240
Trp Gly Lys Gln Val Phe Arg Trp His His Leu Pro Gly Leu His
245 250 255
Trp Leu His His Pro Thr Gly Ala Pro Glu Ile Gly Gly Phe Cys
260 265 270
Gly Val

<210> 23
<211> 188
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506439CD1

<400> 23
Met Leu Gly Lys Leu Ala Met Leu Leu Trp Val Gln Gln Ala Leu
1 5 10 15
Leu Ala Leu Leu Leu Pro Thr Leu Leu Ala Gln Gly Glu Ala Arg
20 25 30
Arg Ser Arg Asn Thr Thr Arg Pro Ala Leu Leu Arg Leu Ser Asp
35 40 45
Tyr Leu Leu Thr Asn Tyr Arg Lys Gly Val Arg Pro Val Arg Asp
50 55 60
Trp Arg Lys Pro Thr Thr Val Ser Ile Asp Val Ile Val Tyr Ala
65 70 75
Ile Leu Asn Val Asp Glu Lys Asn Gln Val Leu Thr Thr Tyr Ile
80 85 90
Trp Tyr Arg Gln Tyr Trp Thr Asp Glu Phe Leu Gln Trp Asn Pro
95 100 105
Glu Asp Phe Asp Asn Ile Thr Lys Leu Ser Ile Pro Thr Asp Ser
110 115 120
Ile Trp Val Pro Asp Ile Leu Ile Asn Glu Phe Val Asp Val Gly

125 130 135
Lys Ser Pro Asn Ile Pro Tyr Val Tyr Ile Arg His Gln His Leu
140 145 150
Phe Val Ala Leu Ala Arg Lys Gly Glu Ile Arg Gln Glu Cys Leu
155 160 165
His Glu Pro Gly Arg Val Gly Val Ala Gly Gly Ala Ala Leu Leu
170 175 180
Ser Gly Val Gln His Gly Lys Gln
185

<210> 24
<211> 111
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509243CD1

<400> 24
Met Pro Ser Ala Gly Leu Cys Ser Cys Trp Gly Gly Arg Val Leu
1 5 10 15
Pro Leu Leu Leu Ala Tyr Val Cys Tyr Leu Leu Leu Gly Ala Thr
20 25 30
Ile Phe Gln Leu Leu Glu Arg Gln Ala Glu Ala Gln Ser Arg Asp
35 40 45
Gln Phe Gln Leu Glu Lys Leu Arg Phe Leu Glu Asn Tyr Thr Cys
50 55 60
Leu Asp Gln Trp Ala Met Glu Gln Phe Val Gln Val Ile Met Glu
65 70 75
Ala Trp Val Lys Gly Val Asn Pro Lys Gly Asn Ser Thr Asn Pro
80 85 90
Ser Asn Trp Asp Phe Gly Ser Ser Phe Phe Phe Ala Gly Thr Val
95 100 105
Val Thr Thr Ile Gly His
110

<210> 25
<211> 46
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509404CD1

<400> 25
Met Lys Phe Leu Leu Thr Thr Ala Phe Leu Ile Leu Ile Ser Leu
1 5 10 15
Trp Val Glu Glu Ala Tyr Ser Lys Glu Lys Ser Ser Lys Lys Gly
20 25 30
Lys Gly Lys Lys Lys Gln Tyr Leu Cys Pro Ser Glu Arg Leu Tyr
35 40 45
His

<210> 26
<211> 204
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509439CD1

<400> 26
Met Phe Ser Arg Ala Gly Val Ala Gly Leu Ser Ala Trp Thr Leu
1 5 10 15
Gln Pro Gln Trp Ile Gln Val Arg Asn Met Ala Thr Leu Lys Asp
20 25 30
Ile Thr Arg Arg Leu Lys Ser Ile Lys Asn Ile Gln Lys Ile Thr
35 40 45
Lys Ser Met Lys Met Val Ala Ala Ala Lys Tyr Ala Arg Ala Glu
50 55 60
Arg Glu Leu Lys Pro Ala Arg Ile Tyr Gly Leu Gly Ser Leu Ala
65 70 75
Leu Tyr Glu Lys Ala Asp Ile Lys Gly Pro Glu Asp Lys Lys Lys
80 85 90
His Leu Leu Ile Gly Val Ser Ser Asp Arg Gly Leu Cys Gly Ala
95 100 105
Ile His Ser Ser Ile Ala Lys Gln Met Lys Ser Glu Val Ala Thr
110 115 120
Leu Thr Ala Ala Gly Lys Glu Val Met Leu Val Gly Ile Gly Asp
125 130 135
Lys Ile Arg Gly Ile Leu Tyr Ser Ser Leu Gln Val Leu Lys Glu
140 145 150
Arg Asn Asp Asp Ser Val Trp Asn Asn Ser Gly Asn His His His
155 160 165
Pro Tyr Pro Lys Asp Leu Ile His Gly Leu Ile Leu Thr Ser Phe
170 175 180
Trp Trp His Ser Lys Lys Trp Glu Glu Ser Pro Pro Leu Leu Glu
185 190 195
Met Arg Gln Ser Leu Pro Leu Asn Tyr
200

<210> 27
<211> 1400
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510202CD1

<400> 27
Met Ser Lys Arg Arg Met Ser Val Gly Gln Gln Thr Trp Ala Leu
1 5 10 15
Leu Cys Lys Asn Cys Leu Lys Lys Trp Arg Met Lys Arg Gln Thr
20 25 30
Leu Leu Glu Trp Leu Phe Ser Phe Leu Leu Val Leu Phe Leu Tyr
35 40 45
Leu Phe Phe Ser Asn Leu His Gln Val His Asp Thr Pro Gln Met

```
                    50                  55                  60
Ser Ser Met Asp Leu Gly Arg Val Asp Ser Phe Asn Asp Thr Asn
                    65                  70                  75
Tyr Val Ile Ala Phe Ala Pro Glu Ser Lys Thr Thr Gln Glu Ile
                    80                  85                  90
Met Asn Lys Val Ala Ser Ala Pro Phe Leu Lys Gly Arg Thr Ile
                    95                  100                 105
Met Gly Trp Pro Asp Glu Lys Ser Met Asp Glu Leu Asp Leu Asn
                    110                 115                 120
Tyr Ser Ile Asp Ala Val Arg Val Ile Phe Thr Asp Thr Phe Ser
                    125                 130                 135
Tyr His Leu Lys Phe Ser Trp Gly His Arg Ile Pro Met Met Lys
                    140                 145                 150
Glu His Arg Asp His Ser Ala His Cys Gln Ala Val Asn Glu Lys
                    155                 160                 165
Met Lys Cys Glu Gly Ser Glu Phe Trp Glu Lys Gly Phe Val Ala
                    170                 175                 180
Phe Gln Ala Ala Ile Asn Ala Ala Ile Ile Glu Ile Ala Thr Asn
                    185                 190                 195
His Ser Val Met Glu Gln Leu Met Ser Val Thr Gly Val His Met
                    200                 205                 210
Lys Ile Leu Pro Phe Val Ala Gln Gly Gly Val Ala Thr Asp Phe
                    215                 220                 225
Phe Ile Phe Phe Cys Ile Ile Ser Phe Ser Thr Phe Ile Tyr Tyr
                    230                 235                 240
Val Ser Val Asn Val Thr Gln Glu Arg Gln Tyr Ile Thr Ser Leu
                    245                 250                 255
Met Thr Met Met Gly Leu Arg Glu Ser Ala Phe Trp Leu Ser Trp
                    260                 265                 270
Gly Leu Met Tyr Ala Gly Phe Ile Leu Ile Met Ala Thr Leu Met
                    275                 280                 285
Ala Leu Ile Val Lys Ser Ala Gln Ile Val Val Leu Thr Gly Phe
                    290                 295                 300
Val Met Val Phe Thr Leu Phe Leu Leu Tyr Gly Leu Ser Leu Ile
                    305                 310                 315
Thr Leu Ala Phe Leu Met Ser Val Leu Ile Lys Lys Pro Phe Leu
                    320                 325                 330
Thr Gly Leu Val Val Phe Leu Leu Ile Val Phe Trp Gly Ile Leu
                    335                 340                 345
Gly Phe Pro Ala Leu Tyr Thr His Leu Pro Ala Phe Leu Glu Trp
                    350                 355                 360
Thr Leu Cys Leu Leu Ser Pro Phe Ala Phe Thr Val Gly Met Ala
                    365                 370                 375
Gln Leu Ile His Leu Asp Tyr Asp Val Asn Ser Asn Ala His Leu
                    380                 385                 390
Asp Ser Ser Gln Asn Pro Tyr Leu Ile Ile Ala Thr Leu Phe Met
                    395                 400                 405
Leu Val Phe Asp Thr Leu Leu Tyr Leu Val Leu Thr Leu Tyr Phe
                    410                 415                 420
Asp Lys Ile Leu Pro Ala Glu Tyr Gly His Arg Cys Ser Pro Leu
                    425                 430                 435
Phe Phe Leu Lys Ser Cys Phe Trp Phe Gln His Gly Arg Ala Asn
                    440                 445                 450
His Val Val Leu Glu Asn Glu Thr Asp Ser Asp Pro Thr Pro Asn
                    455                 460                 465
Asp Cys Phe Glu Pro Val Ser Pro Glu Phe Cys Gly Lys Glu Ala
```

470 475 480
Ile Arg Ile Lys Asn Leu Lys Lys Glu Tyr Ala Gly Lys Cys Glu
485 490 495
Arg Val Glu Ala Leu Lys Gly Val Val Phe Asp Ile Tyr Glu Gly
500 505 510
Gln Ile Thr Ala Leu Leu Gly His Ser Gly Ala Gly Lys Thr Thr
515 520 525
Leu Leu Asn Ile Leu Ser Gly Leu Ser Val Pro Thr Ser Gly Ser
530 535 540
Val Thr Val Tyr Asn His Thr Leu Ser Arg Met Ala Asp Ile Glu
545 550 555
Asn Ile Ser Lys Phe Thr Gly Phe Cys Pro Gln Ser Asn Val Gln
560 565 570
Phe Gly Phe Leu Thr Val Lys Glu Asn Leu Arg Leu Phe Ala Lys
575 580 585
Ile Lys Gly Ile Leu Pro His Glu Val Glu Lys Glu Val Leu Leu
590 595 600
Leu Asp Glu Pro Thr Ala Gly Leu Asp Pro Leu Ser Arg His Arg
605 610 615
Ile Trp Asn Leu Leu Lys Glu Gly Lys Ser Asp Arg Val Ile Leu
620 625 630
Phe Ser Thr Gln Phe Ile Asp Glu Ala Asp Ile Leu Ala Asp Arg
635 640 645
Lys Val Phe Ile Ser Asn Gly Lys Leu Lys Cys Ala Gly Ser Ser
650 655 660
Leu Phe Leu Lys Lys Lys Trp Gly Ile Gly Tyr His Leu Ser Leu
665 670 675
His Leu Asn Glu Arg Cys Asp Pro Glu Ser Ile Thr Ser Leu Val
680 685 690
Lys Gln His Ile Ser Asp Ala Lys Leu Thr Ala Gln Ser Glu Glu
695 700 705
Lys Leu Val Tyr Ile Leu Pro Leu Glu Arg Thr Asn Lys Phe Pro
710 715 720
Glu Leu Tyr Arg Asp Leu Asp Arg Cys Ser Asn Gln Gly Ile Glu
725 730 735
Asp Tyr Gly Val Ser Ile Thr Thr Leu Asn Glu Val Phe Leu Lys
740 745 750
Leu Glu Gly Lys Ser Thr Ile Asp Glu Ser Asp Ile Gly Ile Trp
755 760 765
Gly Gln Leu Gln Thr Asp Gly Ala Lys Asp Ile Gly Ser Leu Val
770 775 780
Glu Leu Glu Gln Val Leu Ser Ser Phe His Glu Thr Arg Lys Thr
785 790 795
Ile Ser Gly Val Ala Leu Trp Arg Gln Gln Val Cys Ala Ile Ala
800 805 810
Lys Val Arg Phe Leu Lys Leu Lys Lys Glu Arg Lys Ser Leu Trp
815 820 825
Thr Ile Leu Leu Leu Phe Gly Ile Ser Phe Ile Pro Gln Leu Leu
830 835 840
Glu His Leu Phe Tyr Glu Ser Tyr Gln Lys Ser Tyr Pro Trp Glu
845 850 855
Leu Ser Pro Asn Thr Tyr Phe Leu Ser Pro Gly Gln Gln Pro Gln
860 865 870
Asp Pro Leu Thr His Leu Leu Val Ile Asn Lys Thr Gly Ser Thr
875 880 885
Ile Asp Asn Phe Leu His Ser Leu Arg Arg Gln Asn Ile Ala Ile

890 895 900
Glu Val Asp Ala Phe Gly Thr Arg Asn Gly Thr Asp Asp Pro Ser
905 910 915
Tyr Asn Gly Ala Ile Ile Val Ser Gly Asp Glu Lys Asp His Arg
920 925 930
Phe Ser Ile Ala Cys Asn Thr Lys Arg Leu Asn Cys Phe Pro Val
935 940 945
Leu Leu Asp Val Ile Ser Asn Gly Leu Leu Gly Ile Phe Asn Ser
950 955 960
Ser Glu His Ile Gln Thr Asp Arg Ser Thr Phe Phe Glu Glu His
965 970 975
Met Asp Tyr Glu Tyr Gly Tyr Arg Ser Asn Thr Phe Phe Trp Ile
980 985 990
Pro Met Ala Ala Ser Phe Thr Pro Tyr Ile Ala Met Ser Ser Ile
995 1000 1005
Gly Asp Tyr Lys Lys Lys Ala His Ser Gln Leu Arg Ile Ser Gly
1010 1015 1020
Leu Tyr Pro Ser Ala Tyr Trp Phe Gly Gln Ala Leu Val Asp Val
1025 1030 1035
Ser Leu Tyr Phe Leu Ile Leu Leu Leu Met Gln Ile Met Asp Tyr
1040 1045 1050
Ile Phe Ser Pro Glu Glu Ile Ile Phe Ile Ile Gln Asn Leu Leu
1055 1060 1065
Ile Gln Ile Leu Cys Ser Ile Gly Tyr Val Ser Ser Pro Val Phe
1070 1075 1080
Leu Thr Tyr Val Ile Ser Phe Ile Phe Arg Asn Gly Arg Lys Asn
1085 1090 1095
Ser Gly Ile Trp Ser Phe Phe Phe Leu Ile Val Val Ile Phe Ser
1100 1105 1110
Ile Val Ala Thr Asp Leu Asn Glu Tyr Gly Phe Leu Gly Leu Phe
1115 1120 1125
Phe Gly Thr Met Leu Ile Pro Pro Phe Thr Leu Ile Gly Ser Leu
1130 1135 1140
Phe Ile Phe Ser Glu Ile Ser Pro Asp Ser Met Asp Tyr Leu Gly
1145 1150 1155
Ala Ser Glu Ser Glu Ile Val Tyr Leu Ala Leu Leu Ile Pro Tyr
1160 1165 1170
Leu His Phe Leu Ile Phe Leu Phe Ile Leu Arg Cys Leu Glu Met
1175 1180 1185
Asn Cys Arg Lys Lys Leu Met Arg Lys Asp Pro Val Phe Arg Ile
1190 1195 1200
Ser Pro Arg Ser Asn Ala Ile Phe Pro Asn Pro Glu Glu Pro Glu
1205 1210 1215
Gly Glu Glu Glu Asp Ile Gln Met Glu Arg Met Arg Thr Val Asn
1220 1225 1230
Ala Met Ala Val Arg Asp Phe Asp Glu Thr Pro Val Ile Ile Ala
1235 1240 1245
Ser Cys Leu Arg Lys Glu Tyr Ala Gly Lys Lys Lys Asn Cys Phe
1250 1255 1260
Ser Lys Arg Lys Lys Thr Ile Ala Thr Arg Asn Val Ser Phe Cys
1265 1270 1275
Val Lys Lys Gly Glu Val Ile Gly Leu Leu Gly His Asn Gly Ala
1280 1285 1290
Gly Lys Ser Thr Thr Ile Lys Met Ile Thr Gly Asp Thr Lys Pro
1295 1300 1305
Thr Ala Gly Gln Val Ile Leu Lys Gly Ser Gly Gly Gly Glu Pro

```
                1310                1315                1320
Leu Gly Phe Leu Gly Tyr Cys Pro Gln Glu Asn Ala Leu Trp Pro
                1325                1330                1335
Asn Leu Thr Val Arg Gln His Leu Glu Val Tyr Ala Ala Val Lys
                1340                1345                1350
Gly Leu Arg Lys Gly Asp Ala Met Ile Ala Ile Thr Arg Leu Val
                1355                1360                1365
Asp Ala Leu Lys Leu Gln Asp Gln Leu Lys Ala Pro Val Lys Thr
                1370                1375                1380
Leu Ser Glu Gly Ile Lys Arg Lys Val Arg Ala Gly Leu Val Val
                1385                1390                1395
Ala Leu Gln Val Pro
                1400
```

<210> 28
<211> 438
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510203CD1

<400> 28

```
Met Gln Ala Thr Arg Asn Ala Ala Asp Trp Trp Leu Ser His Trp
1               5                   10                  15
Ile Ser Gln Leu Lys Ala Glu Asn Ser Ser Gln Glu Ala Gln Pro
                20                  25                  30
Ser Thr Ser Pro Ala Ser Met Gly Leu Phe Ser Pro Gln Leu Leu
                35                  40                  45
Leu Phe Ser Pro Gly Asn Leu Tyr Ile Pro Val Phe Pro Leu Pro
                50                  55                  60
Lys Ala Ala Pro Asn Gly Ser Ser Asp Ile Arg Phe Tyr Leu Thr
                65                  70                  75
Val Tyr Ala Thr Ile Ala Gly Val Asn Ser Leu Cys Thr Leu Leu
                80                  85                  90
Arg Ala Val Leu Phe Ala Ala Gly Thr Leu Gln Ala Ala Ala Thr
                95                  100                 105
Leu His Arg Arg Leu Leu His Arg Val Leu Met Ala Pro Val Thr
                110                 115                 120
Phe Phe Asn Ala Thr Pro Thr Gly Arg Ile Leu Asn Arg Phe Ser
                125                 130                 135
Ser Asp Val Ala Cys Ala Asp Asp Ser Leu Pro Phe Ile Leu Asn
                140                 145                 150
Ile Leu Leu Ala Asn Ala Ala Gly Leu Leu Gly Leu Leu Ala Val
                155                 160                 165
Leu Gly Ser Gly Leu Pro Trp Leu Leu Leu Leu Leu Pro Pro Leu
                170                 175                 180
Ser Ile Met Tyr Tyr His Val Gln Arg His Tyr Arg Ala Ser Ser
                185                 190                 195
Arg Glu Leu Arg Arg Leu Gly Ser Leu Thr Leu Ser Pro Leu Tyr
                200                 205                 210
Ser His Leu Ala Asp Thr Leu Ala Gly Leu Ser Val Leu Arg Ala
                215                 220                 225
Thr Gly Ala Thr Tyr Arg Phe Glu Glu Glu Asn Leu Arg Leu Leu
                230                 235                 240
```

Glu Leu Asn Gln Arg Cys Gln Phe Ala Thr Ser Ala Thr Met Gln
245 250 255
Trp Leu Asp Ile Arg Leu Gln Leu Met Gly Ala Ala Val Val Ser
260 265 270
Ala Ile Ala Gly Ile Ala Leu Val Gln His Gln Gln Gly Leu Ala
275 280 285
Asn Pro Gly Leu Val Gly Leu Ser Leu Ser Tyr Ala Leu Ser Leu
290 295 300
Thr Gly Leu Leu Ser Gly Leu Val Ser Ser Phe Thr Gln Thr Glu
305 310 315
Ala Met Leu Val Ser Val Glu Arg Leu Glu Glu Tyr Thr Cys Asp
320 325 330
Leu Pro Gln Glu Pro Gln Gly Gln Pro Leu Gln Val Gly Leu Tyr
335 340 345
Pro His Pro Arg Pro Lys Leu Trp Asn Pro Glu Gly Pro Ser Leu
350 355 360
Pro His Asn Ser Phe Leu Phe Ala His Pro Ser Phe Ser Ala Pro
365 370 375
Ile Thr Ser Leu His Asp Asp His Asn Ser Ser Pro Cys Pro Phe
380 385 390
Phe Pro Ile Ser His Ser Leu Ile Pro Leu Thr Leu Ser Ile Ser
395 400 405
His Tyr Ser Pro Leu Leu Thr Ile Ala Pro His Leu Pro Tyr Leu
410 415 420
Pro Phe Pro Val Cys Leu Pro Pro Met Asp Pro Thr Ser Trp Ala
425 430 435
Pro Ala Gly

<210> 29
<211> 871
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510208CD1

<400> 29
Met Gly Phe Leu His Gln Leu Gln Leu Leu Leu Trp Lys Asn Val
1 5 10 15
Thr Leu Lys Arg Arg Ser Pro Trp Val Leu Ala Phe Glu Ile Phe
20 25 30
Ile Pro Leu Val Leu Phe Phe Ile Leu Leu Gly Leu Arg Gln Lys
35 40 45
Lys Pro Thr Ile Ser Val Lys Glu Val Ser Phe Tyr Thr Ala Ala
50 55 60
Pro Leu Thr Ser Ala Gly Ile Leu Pro Val Met Gln Ser Leu Cys
65 70 75
Pro Asp Gly Gln Arg Asp Glu Phe Gly Phe Leu Gln Tyr Ala Asn
80 85 90
Ser Thr Val Thr Gln Leu Leu Glu Arg Leu Asp Arg Val Val Glu
95 100 105
Glu Gly Asn Leu Phe Asp Pro Ala Arg Pro Ser Leu Gly Ser Glu
110 115 120
Leu Glu Ala Leu Arg Gln His Leu Glu Ala Leu Ser Ala Gly Pro

125 130 135
Gly Thr Ser Gly Ser His Leu Asp Arg Ser Thr Val Ser Ser Phe
140 145 150
Ser Leu Asp Ser Val Ala Arg Asn Pro Gln Glu Leu Trp Arg Phe
155 160 165
Leu Thr Gln Asn Leu Ser Leu Pro Asn Ser Thr Ala Gln Ala Leu
170 175 180
Leu Ala Ala Arg Val Asp Pro Pro Glu Val Tyr His Leu Leu Phe
185 190 195
Gly Pro Ser Ser Ala Leu Asp Ser Gln Ser Gly Leu His Lys Gly
200 205 210
Gln Glu Pro Trp Ser Arg Leu Gly Gly Asn Pro Leu Phe Arg Met
215 220 225
Glu Glu Leu Leu Leu Ala Pro Ala Leu Leu Glu Gln Leu Thr Cys
230 235 240
Thr Pro Gly Ser Gly Glu Leu Gly Arg Ile Leu Thr Val Pro Glu
245 250 255
Ser Gln Lys Gly Ala Leu Gln Gly Tyr Arg Asp Ala Val Cys Ser
260 265 270
Gly Gln Ala Ala Ala Arg Ala Arg Arg Phe Ser Gly Leu Ser Ala
275 280 285
Glu Leu Arg Asn Gln Leu Asp Val Ala Lys Val Ser Gln Gln Leu
290 295 300
Gly Leu Asp Ala Pro Asn Gly Ser Asp Ser Ser Pro Gln Ala Pro
305 310 315
Pro Pro Arg Arg Leu Gln Ala Leu Leu Gly Asp Leu Leu Asp Ala
320 325 330
Gln Lys Val Leu Gln Asp Val Asp Val Leu Ser Ala Leu Ala Leu
335 340 345
Leu Leu Pro Gln Gly Ala Cys Thr Gly Arg Thr Pro Gly Pro Pro
350 355 360
Ala Ser Gly Ala Gly Gly Ala Ala Asn Gly Thr Gly Ala Gly Ala
365 370 375
Val Met Gly Pro Asn Ala Thr Ala Glu Glu Gly Ala Pro Ser Ala
380 385 390
Ala Ala Leu Ala Thr Pro Asp Thr Leu Gln Gly Gln Cys Ser Ala
395 400 405
Phe Val Gln Leu Trp Ala Gly Leu Gln Pro Ile Leu Cys Gly Asn
410 415 420
Asn Arg Thr Ile Glu Pro Glu Ala Leu Arg Arg Gly Asn Met Ser
425 430 435
Ser Leu Gly Phe Thr Ser Lys Glu Gln Arg Asn Leu Gly Leu Leu
440 445 450
Val His Leu Met Thr Ser Asn Pro Lys Ile Leu Tyr Ala Pro Ala
455 460 465
Gly Ser Glu Val Asp Arg Val Ile Leu Lys Ala Asn Glu Thr Phe
470 475 480
Ala Phe Val Gly Asn Val Thr His Tyr Ala Gln Val Trp Leu Asn
485 490 495
Ile Ser Ala Glu Ile Arg Ser Phe Leu Glu Gln Gly Arg Leu Gln
500 505 510
Gln His Leu Arg Trp Leu Gln Gln Tyr Val Ala Glu Leu Arg Leu
515 520 525
His Pro Glu Ala Leu Asn Leu Ser Leu Asp Glu Leu Pro Pro Ala
530 535 540
Leu Arg Gln Asp Asn Phe Ser Leu Pro Ser Gly Met Ala Leu Leu

```
                545                 550                 555
Gln Gln Leu Asp Thr Ile Asp Asn Ala Ala Cys Gly Trp Ile Gln
                560                 565                 570
Phe Met Ser Lys Val Ser Val Asp Ile Phe Lys Gly Phe Pro Asp
                575                 580                 585
Glu Glu Ser Ile Val Asn Tyr Thr Leu Asn Gln Ala Tyr Gln Asp
                590                 595                 600
Asn Val Thr Val Phe Ala Ser Val Ile Phe Gln Thr Arg Lys Asp
                605                 610                 615
Gly Ser Leu Pro Pro His Val His Tyr Lys Ile Arg Gln Asn Ser
                620                 625                 630
Ser Phe Thr Glu Lys Thr Asn Glu Ile Arg Arg Ala Tyr Trp Arg
                635                 640                 645
Pro Gly Pro Asn Thr Gly Gly Arg Phe Tyr Phe Leu Tyr Gly Phe
                650                 655                 660
Val Trp Ile Gln Asp Met Met Glu Arg Ala Ile Ile Asp Thr Phe
                665                 670                 675
Val Gly His Asp Val Val Glu Pro Gly Ser Tyr Val Gln Met Phe
                680                 685                 690
Pro Tyr Pro Cys Tyr Thr Arg Asp Asp Phe Leu Phe Val Ile Glu
                695                 700                 705
His Met Met Pro Leu Cys Met Val Ile Ser Trp Val Tyr Ser Val
                710                 715                 720
Ala Met Thr Ile Gln His Ile Val Ala Glu Lys Glu His Arg Leu
                725                 730                 735
Lys Glu Val Arg Gly Pro Gly Leu Ser Leu Glu Ala Arg Ala Gly
                740                 745                 750
Arg Glu Gly Arg Arg Pro Pro Arg Gly Leu Pro Gln Ala Pro Gly
                755                 760                 765
Pro Pro Ala Gly Asp Glu Asp His Gly Pro Glu Gln Arg Gly Ala
                770                 775                 780
Leu Gly Gly Leu Val His His Arg Leu Cys Ala Ala Val His Leu
                785                 790                 795
Arg Asp Ser Thr His Arg His Pro Glu Val Arg Pro Gly Ala Tyr
                800                 805                 810
Ala Gln Pro Arg Gly His His Leu Ala Leu Pro Gly Ser Leu Arg
                815                 820                 825
Gly Gly His His His Val Leu Leu Pro Gly Val Cys Ala Val Leu
                830                 835                 840
Gln Gly Gln Ala Gly Leu Gly Leu Arg Trp His His Leu Leu Pro
                845                 850                 855
Glu Leu Arg Ala Leu His Val Arg Gly Asp Pro Arg Gly Gly Gly
                860                 865                 870
Ala
```

<210> 30
<211> 104
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510446CD1

<400> 30

Met Glu Gln Ser Arg Ser Gln Gln Arg Gly Gly Glu Gln Ser Trp
1 5 10 15
Trp Gly Ser Asp Pro Gln Tyr Gln Tyr Met Pro Phe Glu His Cys
20 25 30
Thr Ser Tyr Gly Leu Pro Ser Glu Asn Gly Gly Leu Gln His Arg
35 40 45
Leu Arg Lys Asp Ala Gly Pro Arg His Asn Val His Pro Thr Gln
50 55 60
Ile Tyr Gly His His Lys Glu Gln Phe Ser Asp Arg Glu Gln Asp
65 70 75
Ile Gly Met Pro Lys Lys Thr Gly Ser Ser Ser Thr Val Asp Ser
80 85 90
Lys Asp Glu Asp His Tyr Ser Lys Cys Gln Gly Asp Gly Asp
95 100

<210> 31
<211> 336
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505294CD1

<400> 31
Met Ala Ser Asp Pro Ile Phe Thr Leu Ala Pro Pro Leu His Cys
1 5 10 15
His Tyr Gly Ala Phe Pro Pro Asn Ala Ser Gly Trp Glu Gln Pro
20 25 30
Pro Asn Ala Ser Gly Val Ser Val Ala Ser Ala Ala Leu Ala Ala
35 40 45
Ser Ala Ala Ser Arg Val Ala Thr Ser Thr Asp Pro Ser Cys Ser
50 55 60
Gly Phe Ala Pro Pro Asp Phe Asn His Cys Leu Lys Asp Trp Asp
65 70 75
Tyr Asn Gly Leu Pro Val Leu Thr Thr Asn Ala Ile Gly Gln Trp
80 85 90
Asp Leu Val Cys Asp Leu Gly Trp Gln Val Ile Leu Glu Gln Ile
95 100 105
Leu Phe Ile Leu Gly Phe Ala Ser Gly Tyr Leu Phe Leu Gly Tyr
110 115 120
Pro Ala Asp Arg Phe Gly Arg Arg Gly Ile Val Leu Leu Thr Leu
125 130 135
Gly Leu Val Gly Pro Cys Gly Val Gly Gly Ala Ala Ala Gly Ser
140 145 150
Ser Thr Gly Val Met Ala Leu Arg Phe Leu Leu Gly Phe Leu Leu
155 160 165
Ala Gly Val Asp Leu Gly Val Tyr Leu Met Arg Leu Glu Leu Cys
170 175 180
Asp Pro Thr Gln Arg Leu Arg Val Ala Leu Ala Gly Glu Leu Val
185 190 195
Gly Val Gly Gly His Phe Leu Phe Leu Gly Leu Ala Leu Val Ser
200 205 210
Lys Asp Trp Arg Phe Leu Gln Arg Met Ile Thr Ala Pro Cys Ile
215 220 225
Leu Phe Leu Phe Tyr Gly Trp Pro Gly Leu Phe Leu Glu Ser Ala

```
                230                     235                     240
Arg Trp Leu Ile Val Lys Arg Gln Ile Glu Glu Ala Gln Ser Val
                245                     250                     255
Leu Arg Ile Leu Ala Glu Arg Asn Arg Pro His Gly Gln Met Leu
                260                     265                     270
Gly Glu Glu Ala Gln Glu Ala Leu Gln Ala Ser Leu Pro Met Pro
                275                     280                     285
Phe Ala Thr Ala Thr Ser Leu Trp Glu Glu Glu Gly Ala His Arg
                290                     295                     300
Thr Ser Thr Cys Ala Leu Cys Trp Pro Ala Ala Pro Gln Pro Trp
                305                     310                     315
Pro Val Ser Ser Trp Gly Ser Pro Trp Thr Asp Leu Ala Ala Gly
                320                     325                     330
Ala Ser Phe Phe Ser Pro
                335

<210> 32
<211> 271
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505631CD1

<400> 32
Met Asp Asp Phe Ile Ser Ile Ser Leu Leu Ser Leu Ala Met Leu
  1               5                  10                  15
Val Gly Cys Tyr Val Ala Gly Ile Ile Pro Leu Ala Val Asn Phe
                 20                  25                  30
Ser Glu Glu Arg Leu Lys Leu Val Thr Val Leu Gly Ala Gly Leu
                 35                  40                  45
Leu Cys Gly Thr Ala Leu Ala Val Ile Val Pro Glu Gly Val His
                 50                  55                  60
Ala Leu Tyr Glu Asp Ile Leu Glu Gly Lys His His Gln Ala Ser
                 65                  70                  75
Glu Thr His Asn Val Ile Ala Ser Asp Lys Ala Ala Glu Lys Ser
                 80                  85                  90
Val Val His Glu His Glu His Ser His Asp His Thr Gln Leu His
                 95                 100                 105
Ala Tyr Ile Gly Val Ser Leu Val Leu Gly Phe Val Phe Met Leu
                110                 115                 120
Leu Val Asp Gln Ile Gly Asn Ser His Val His Ser Thr Asp Asp
                125                 130                 135
Pro Glu Ala Ala Arg Ser Ser Asn Ser Lys Ile Thr Thr Thr Leu
                140                 145                 150
Gly Leu Val Val His Ala Ala Ala Asp Gly Val Ala Leu Gly Ala
                155                 160                 165
Ala Ala Ser Thr Ser Gln Thr Ser Val Gln Leu Ile Val Phe Val
                170                 175                 180
Ala Ile Met Leu His Lys Ala Pro Ala Ala Phe Gly Leu Val Ser
                185                 190                 195
Phe Leu Met His Ala Gly Leu Glu Arg Asn Arg Ile Arg Lys His
                200                 205                 210
Leu Leu Val Phe Ala Leu Ala Ala Pro Val Met Ser Met Val Thr
                215                 220                 225
```

Tyr Leu Gly Leu Ser Lys Ser Ser Lys Glu Ala Leu Ser Glu Val
                230                 235                 240
Asn Ala Thr Gly Val Ala Met Leu Phe Ser Ala Gly Thr Phe Leu
                245                 250                 255
Tyr Val Ala Thr Val Arg Lys Val Ala Gln Ile Gly Tyr Ser Cys
                260                 265                 270
Met

<210> 33
<211> 107
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506561CD1

<400> 33
Met Ala Trp Gln Met Met Gln Leu Leu Leu Leu Ala Leu Val Thr
1               5                   10                  15
Ala Ala Gly Ser Ala Gln Pro Arg Ser Ala Arg Ala Arg Thr Asp
                20                  25                  30
Leu Leu Asn Val Cys Met Asn Ala Lys His His Lys Thr Gln Pro
                35                  40                  45
Ser Pro Glu Asp Glu Leu Tyr Gly Gln Cys Ser Pro Trp Lys Lys
                50                  55                  60
Asn Ala Cys Cys Thr Ala Ser Thr Ser Gln Glu Leu His Lys Asp
                65                  70                  75
Thr Ser Arg Leu Tyr Asn Phe Asn Trp Asp His Cys Gly Gln Pro
                80                  85                  90
Glu Leu Ala Gln Arg Ala His Ser Glu Arg Ala Pro Val Gln Arg
                95                  100                 105
Gly Leu

<210> 34
<211> 249
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510733CD1

<400> 34
Met Gln Pro Glu Gly Ala Glu Lys Gly Lys Ser Phe Lys Gln Arg
1               5                   10                  15
Leu Val Leu Lys Ser Ser Leu Ala Lys Glu Thr Leu Ser Glu Phe
                20                  25                  30
Leu Gly Thr Phe Ile Leu Ile Val Leu Gly Cys Gly Cys Val Ala
                35                  40                  45
Gln Ala Ile Leu Ser Arg Gly Arg Phe Gly Gly Val Ile Thr Ile
                50                  55                  60
Asn Val Gly Phe Ser Met Ala Val Ala Met Ala Ile Tyr Val Ala
                65                  70                  75

Gly Gly Val Ser Asp Gly Leu Met Ser Phe Ala Gly Gly Lys Leu
80 85 90
Leu Ile Val Gly Glu Asn Ala Thr Ala His Ile Phe Ala Thr Tyr
95 100 105
Pro Ala Pro Tyr Leu Ser Leu Ala Asn Ala Phe Ala Asp Gln Val
110 115 120
Val Ala Thr Met Ile Leu Leu Ile Ile Val Phe Ala Ile Phe Asp
125 130 135
Ser Arg Asn Leu Gly Ala Pro Arg Gly Leu Glu Pro Ile Ala Ile
140 145 150
Gly Leu Leu Ile Ile Val Ile Ala Ser Ser Leu Gly Leu Asn Ser
155 160 165
Gly Cys Ala Met Asn Pro Ala Arg Asp Leu Ser Pro Arg Leu Phe
170 175 180
Thr Ala Leu Ala Gly Trp Gly Phe Glu Val Phe Arg Ala Gly Asn
185 190 195
Asn Phe Trp Trp Ile Pro Val Val Gly Pro Leu Val Gly Ala Val
200 205 210
Ile Gly Gly Leu Ile Tyr Val Leu Val Ile Glu Ile His His Pro
215 220 225
Glu Pro Asp Ser Val Phe Lys Ala Glu Gln Ser Glu Asp Lys Pro
230 235 240
Glu Lys Tyr Glu Leu Ser Val Ile Met
245

<210> 35
<211> 216
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510734CD1

<400> 35
Met Gln Pro Glu Gly Ala Glu Lys Gly Lys Ser Phe Lys Gln Arg
1 5 10 15
Leu Val Leu Lys Ser Ser Leu Ala Lys Glu Thr Leu Ser Glu Phe
20 25 30
Leu Gly Thr Phe Ile Leu Ile Val Leu Gly Cys Gly Cys Val Ala
35 40 45
Gln Ala Ile Leu Ser Arg Gly Arg Phe Gly Gly Val Ile Thr Ile
50 55 60
Asn Val Gly Phe Ser Met Ala Val Ala Met Ala Ile Tyr Val Ala
65 70 75
Gly Gly Val Ser Gly Gly His Ile Asn Pro Ala Val Ser Leu Ala
80 85 90
Met Cys Leu Phe Gly Arg Met Lys Trp Phe Lys Leu Pro Phe Tyr
95 100 105
Val Gly Ala Gln Phe Leu Gly Ala Phe Val Gly Ala Ala Thr Val
110 115 120
Phe Gly Ile Tyr Tyr Asp Gly Leu Met Ser Phe Ala Gly Gly Lys
125 130 135
Leu Leu Ile Val Gly Glu Asn Ala Thr Ala His Ile Phe Ala Thr
140 145 150
Tyr Pro Ala Pro Tyr Leu Ser Leu Ala Asn Ala Phe Ala Asp Gln

155 160 165
Lys Leu Gly Ser Pro Gln Arg Pro Arg Ala His Cys His Arg Pro
170 175 180
Pro Asp Tyr Cys His Cys Phe Leu Pro Gly Thr Glu Gln Trp Leu
185 190 195
Cys His Glu Pro Ser Ser Arg Pro Glu Ser Gln Thr Phe His Cys
200 205 210
Leu Gly Arg Leu Gly Val
215

<210> 36
<211> 223
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503977CD1

<400> 36
Met Ala Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala
1 5 10 15
Pro Trp Gly Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys
20 25 30
Gly Ser Phe Pro Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp
35 40 45
Gly Val Gln Phe Pro Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu
50 55 60
Val Asp Asp Gly Thr His Gly Cys Leu Gly Gly Glu Asn Arg Phe
65 70 75
Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys Thr Gly Val
80 85 90
Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Leu Ile Asp
95 100 105
Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln Ala
110 115 120
Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp
125 130 135
Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly
140 145 150
Gly Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe
155 160 165
Pro Lys Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile
170 175 180
Met Thr Arg Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly
185 190 195
Ser Glu Glu Phe Glu Thr Ile Val Leu Lys Ala Leu Val Lys Val
200 205 210
Leu Pro Ser Arg Ser Phe Pro His Gly Arg Pro Ala Glu
215 220

<210> 37
<211> 394
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505084CD1

<400> 37
Met Glu Ser Gly Thr Ser Ser Pro Gln Pro Pro Gln Leu Asp Pro
1 5 10 15
Leu Asp Ala Phe Pro Gln Lys Gly Leu Glu Pro Gly Asp Ile Ala
20 25 30
Val Leu Val Leu Tyr Phe Leu Phe Val Leu Ala Val Gly Leu Trp
35 40 45
Ser Thr Val Lys Thr Lys Arg Asp Thr Val Lys Gly Tyr Phe Leu
50 55 60
Ala Gly Gly Asp Met Val Trp Trp Pro Val Gly Ala Ser Leu Phe
65 70 75
Ala Ser Asn Val Gly Ser Gly His Phe Ile Gly Leu Ala Gly Ser
80 85 90
Gly Ala Ala Thr Gly Ile Ser Val Ser Ala Tyr Glu Leu Asn Gly
95 100 105
Leu Phe Ser Val Leu Met Leu Ala Trp Ile Phe Leu Pro Ile Tyr
110 115 120
Ile Ala Gly Gln Val Thr Thr Met Pro Glu Tyr Leu Arg Lys Arg
125 130 135
Phe Gly Gly Ile Arg Ile Pro Ile Ile Leu Ala Val Leu Tyr Leu
140 145 150
Phe Ile Tyr Ile Phe Thr Lys Ile Ser Val Asp Met Tyr Ala Gly
155 160 165
Ala Ile Phe Ile Gln Gln Ser Leu His Leu Asp Leu Tyr Leu Ala
170 175 180
Ile Val Gly Leu Leu Ala Ile Thr Ala Val Tyr Thr Val Ala Gly
185 190 195
Gly Leu Ala Ala Val Ile Tyr Thr Asp Ala Leu Gln Thr Leu Ile
200 205 210
Met Leu Ile Gly Ala Leu Thr Leu Met Gly Tyr Ser Phe Ala Ala
215 220 225
Val Gly Gly Met Glu Gly Leu Lys Glu Lys Tyr Phe Leu Ala Leu
230 235 240
Ala Ser Asn Arg Ser Glu Asn Ser Ser Cys Gly Leu Pro Arg Glu
245 250 255
Asp Ala Phe His Ile Phe Arg Asp Pro Leu Thr Ser Asp Leu Pro
260 265 270
Trp Pro Gly Val Leu Phe Gly Met Ser Ile Pro Ser Leu Trp Tyr
275 280 285
Trp Cys Thr Asp Gln Val Ile Val Gln Arg Thr Leu Ala Ala Lys
290 295 300
Asn Leu Ser His Ala Lys Gly Gly Ala Leu Met Ala Ala Tyr Leu
305 310 315
Lys Val Leu Pro Leu Phe Ile Met Val Phe Pro Gly Met Val Ser
320 325 330
Arg Ile Leu Phe Pro Asp Gln Val Ala Cys Ala Asp Pro Glu Ile
335 340 345
Cys Gln Lys Ile Cys Ser Asn Pro Ser Gly Cys Ser Asp Ile Ala
350 355 360
Tyr Pro Lys Leu Val Leu Glu Leu Leu Pro Thr Val Pro Ala Pro
365 370 375
Ser Ser Pro Trp Thr Ser Gly Ile Thr Ser Gly Leu Gly His Leu

```
                380                 385                 390
Arg Arg Ser Ser


<210> 38
<211> 202
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506950CD1

<400> 38
Met Lys Thr Lys Leu Asn Ile Tyr Asn Met Gln Phe Leu Leu Phe
  1               5                  10                  15
Val Phe Leu Val Trp Asp Pro Ala Arg Leu Val Leu Ala Asn Ile
                 20                  25                  30
Gln Glu Asp Glu Ala Lys Asn Asn Ile Thr Ile Phe Thr Arg Ile
                 35                  40                  45
Leu Asp Arg Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly
                 50                  55                  60
Leu Gly Asp Ser Ile Thr Glu Val Phe Thr Asn Ile Tyr Val Thr
                 65                  70                  75
Ser Phe Gly Pro Val Ser Asp Thr Asp Met Glu Tyr Thr Ile Asp
                 80                  85                  90
Val Phe Phe Arg Gln Lys Trp Lys Asp Glu Arg Leu Lys Phe Lys
                 95                 100                 105
Gly Pro Met Asn Ile Leu Arg Leu Asn Asn Leu Met Ala Ser Lys
                110                 115                 120
Ile Trp Thr Pro Asp Thr Phe Phe His Asn Gly Lys Lys Ser Val
                125                 130                 135
Ala His Asn Met Thr Met Pro Asn Lys Leu Leu Arg Ile Gln Asp
                140                 145                 150
Asp Gly Thr Leu Leu Tyr Thr Met Arg Ser Asn Asn Cys Pro Asn
                155                 160                 165
Asn Asp Asn Ser Lys His Gln Cys Ser Glu Phe Ser Pro Gln Ser
                170                 175                 180
Gly Leu Cys Asn Cys His Gly Leu Val Tyr Cys Cys Leu Leu Cys
                185                 190                 195
Ile Cys Val Leu Cys Pro Asn
                200

<210> 39
<211> 337
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506951CD1

<400> 39
Met Lys Thr Lys Leu Asn Ile Tyr Asn Met Gln Phe Leu Leu Phe
  1               5                  10                  15
Val Phe Leu Val Trp Asp Pro Ala Arg Leu Val Leu Ala Asn Ile
```

```
                20                  25                  30
Gln Glu Asp Glu Ala Lys Asn Asn Ile Thr Ile Phe Thr Arg Ile
                35                  40                  45
Leu Asp Arg Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly
                50                  55                  60
Leu Gly Asp Ser Ile Thr Glu Val Phe Thr Asn Ile Tyr Val Thr
                65                  70                  75
Ser Phe Gly Pro Val Ser Asp Thr Asp Met Glu Tyr Thr Ile Asp
                80                  85                  90
Val Phe Phe Arg Gln Lys Trp Lys Asp Glu Arg Leu Lys Phe Lys
                95                 100                 105
Gly Pro Met Asn Ile Leu Arg Leu Asn Asn Leu Met Ala Ser Lys
               110                 115                 120
Ile Trp Thr Pro Asp Thr Phe Phe His Asn Gly Lys Lys Ser Val
               125                 130                 135
Ala His Asn Met Thr Met Pro Asn Lys Leu Leu Arg Ile Gln Asp
               140                 145                 150
Asp Gly Thr Leu Leu Tyr Thr Met Arg Leu Thr Val Gln Ala Glu
               155                 160                 165
Cys Pro Met His Leu Glu Asp Phe Pro Met Asp Ala His Ser Cys
               170                 175                 180
Pro Leu Lys Phe Gly Ser Tyr Ala Tyr Thr Thr Ser Glu Val Thr
               185                 190                 195
Tyr Ile Trp Thr Tyr Asn Ala Ser Asp Ser Val Gln Val Ala Pro
               200                 205                 210
Asp Gly Ser Arg Leu Asn Gln Tyr Asp Leu Leu Gly Gln Ser Ile
               215                 220                 225
Gly Lys Glu Thr Ile Lys Ser Ser Thr Gly Glu Tyr Thr Val Met
               230                 235                 240
Thr Ala His Phe His Leu Lys Arg Lys Ile Gly Tyr Phe Val Ile
               245                 250                 255
Gln Thr Tyr Leu Pro Cys Ile Met Thr Val Ile Leu Ser Gln Val
               260                 265                 270
Ser Phe Trp Leu Asn Arg Glu Ser Val Pro Ala Arg Thr Val Phe
               275                 280                 285
Glu Lys Arg Lys Gly Phe Arg Tyr Asp Thr Glu Gln Arg Leu Cys
               290                 295                 300
Ser Gly Cys Cys Gln Leu Cys Pro Glu Ser Phe Lys Arg Ser Ser
               305                 310                 315
Ser Leu His His Leu Gln Glu Cys Asn His Ala Arg Thr Gln Gln
               320                 325                 330
Glu Ala Arg Lys Gln Ala Ser
               335
```

<210> 40
<211> 114
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506954CD1

<400> 40

```
Met Lys Thr Lys Leu Asn Ile Tyr Asn Met Gln Phe Leu Leu Phe
  1               5                  10                  15
```

Val Phe Leu Val Trp Asp Pro Ala Arg Leu Val Leu Ala Asn Ile
20 25 30
Gln Glu Asp Glu Ala Lys Asn Asn Ile Thr Ile Phe Thr Arg Ile
35 40 45
Leu Asp Arg Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly
50 55 60
Leu Gly Glu Lys Arg Lys Gly Phe Arg Tyr Asp Thr Glu Gln Arg
65 70 75
Leu Cys Ser Gly Cys Cys Gln Leu Cys Pro Glu Ser Phe Lys Arg
80 85 90
Ser Ser Ser Leu His His Leu Gln Glu Cys Asn His Ala Arg Thr
95 100 105
Gln Gln Glu Ala Arg Lys Gln Ala Ser
110

<210> 41
<211> 400
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506956CD1

<400> 41
Met Lys Thr Lys Leu Asn Ile Tyr Asn Met Gln Phe Leu Leu Phe
1 5 10 15
Val Phe Leu Val Trp Asp Pro Ala Arg Leu Val Leu Ala Asn Ile
20 25 30
Gln Glu Asp Glu Ala Lys Asn Asn Ile Thr Ile Phe Thr Arg Ile
35 40 45
Leu Asp Arg Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly
50 55 60
Leu Gly Asp Ser Ile Thr Glu Val Phe Thr Asn Ile Tyr Val Thr
65 70 75
Ser Phe Gly Pro Val Ser Asp Thr Asp Met Glu Tyr Thr Ile Asp
80 85 90
Val Phe Phe Arg Gln Lys Trp Lys Asp Glu Arg Leu Lys Phe Lys
95 100 105
Gly Pro Met Asn Ile Leu Arg Leu Asn Asn Leu Met Ala Ser Lys
110 115 120
Ile Trp Thr Pro Asp Thr Phe Phe His Asn Gly Lys Lys Ser Val
125 130 135
Ala His Asn Met Thr Met Pro Asn Lys Leu Leu Arg Ile Gln Asp
140 145 150
Asp Gly Thr Leu Leu Tyr Thr Met Arg Leu Thr Val Gln Ala Glu
155 160 165
Cys Pro Met His Leu Glu Asp Phe Pro Met Asp Ala His Ser Cys
170 175 180
Pro Leu Lys Phe Gly Ser Tyr Ala Tyr Thr Thr Ser Glu Val Thr
185 190 195
Tyr Ile Trp Thr Tyr Asn Ala Ser Asp Ser Val Gln Val Ala Pro
200 205 210
Asp Gly Ser Arg Leu Asn Gln Tyr Asp Leu Leu Gly Gln Ser Ile
215 220 225
Gly Lys Glu Thr Ile Lys Ser Ser Thr Gly Val Thr Thr Val Leu

```
                230                 235                 240
Thr Met Thr Thr Leu Ser Ile Ser Ala Arg Asn Ser Leu Pro Lys
                245                 250                 255
Val Ala Tyr Ala Thr Ala Met Asp Trp Phe Ile Ala Val Cys Tyr
                260                 265                 270
Ala Phe Val Phe Ser Ala Leu Ile Glu Phe Ala Thr Val Asn Tyr
                275                 280                 285
Phe Thr Lys Arg Gly Trp Ala Trp Asp Gly Lys Ser Val Val Asn
                290                 295                 300
Asp Lys Lys Lys Glu Lys Ala Ser Val Met Ile Gln Asn Asn Ala
                305                 310                 315
Tyr Ala Val Ala Val Ala Asn Tyr Ala Pro Asn Leu Ser Lys Asp
                320                 325                 330
Pro Val Leu Ser Thr Ile Ser Lys Ser Ala Thr Thr Pro Glu Pro
                335                 340                 345
Asn Lys Lys Pro Glu Asn Lys Pro Ala Glu Ala Lys Lys Thr Phe
                350                 355                 360
Asn Ser Val Ser Lys Ile Asp Arg Met Ser Arg Ile Val Phe Pro
                365                 370                 375
Val Leu Phe Gly Thr Phe Asn Leu Val Tyr Trp Ala Thr Tyr Leu
                380                 385                 390
Asn Arg Glu Pro Val Leu Gly Val Ser Pro
                395                 400
```

```
<210> 42
<211> 403
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506959CD1

<400> 42
Met Lys Thr Lys Leu Asn Ile Tyr Asn Met Gln Phe Leu Leu Phe
 1               5                  10                  15
Val Phe Leu Val Trp Asp Pro Ala Arg Leu Val Leu Ala Asn Ile
                20                  25                  30
Gln Glu Asp Glu Ala Lys Asn Asn Ile Thr Ile Phe Thr Arg Ile
                35                  40                  45
Leu Asp Arg Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly
                50                  55                  60
Leu Gly Asp Ser Ile Thr Glu Val Phe Thr Asn Ile Tyr Val Thr
                65                  70                  75
Ser Phe Gly Pro Val Ser Asp Thr Asp Met Glu Tyr Thr Ile Asp
                80                  85                  90
Val Phe Phe Arg Gln Lys Trp Lys Asp Glu Arg Leu Lys Phe Lys
                95                  100                 105
Gly Pro Met Asn Ile Leu Arg Leu Asn Asn Leu Met Ala Ser Lys
                110                 115                 120
Ile Trp Thr Pro Asp Thr Phe Phe His Asn Gly Lys Lys Ser Val
                125                 130                 135
Ala His Asn Met Thr Met Pro Asn Lys Leu Leu Arg Ile Gln Asp
                140                 145                 150
Asp Gly Thr Leu Leu Tyr Thr Met Arg Leu Thr Val Gln Ala Glu
                155                 160                 165
```

Cys Pro Met His Leu Glu Asp Phe Pro Met Asp Ala His Ser Cys
170 175 180
Pro Leu Lys Phe Gly Ser Cys Glu Tyr Thr Val Met Thr Ala His
185 190 195
Phe His Leu Lys Arg Lys Ile Gly Tyr Phe Val Ile Gln Thr Tyr
200 205 210
Leu Pro Cys Ile Met Thr Val Ile Leu Ser Gln Val Ser Phe Trp
215 220 225
Leu Asn Arg Glu Ser Val Pro Ala Arg Thr Val Phe Gly Val Thr
230 235 240
Thr Val Leu Thr Met Thr Thr Leu Ser Ile Ser Ala Arg Asn Ser
245 250 255
Leu Pro Lys Val Ala Tyr Ala Thr Ala Met Asp Trp Phe Ile Ala
260 265 270
Val Cys Tyr Ala Phe Val Phe Ser Ala Leu Ile Glu Phe Ala Thr
275 280 285
Val Asn Tyr Phe Thr Lys Arg Gly Trp Ala Trp Asp Gly Lys Ser
290 295 300
Val Val Asn Asp Lys Lys Lys Glu Lys Ala Ser Val Met Ile Gln
305 310 315
Asn Asn Ala Tyr Ala Val Ala Val Ala Asn Tyr Ala Pro Asn Leu
320 325 330
Ser Lys Asp Pro Val Leu Ser Thr Ile Ser Lys Ser Ala Thr Thr
335 340 345
Pro Glu Pro Asn Lys Lys Pro Glu Asn Lys Pro Ala Glu Ala Lys
350 355 360
Lys Thr Phe Asn Ser Val Ser Lys Ile Asp Arg Met Ser Arg Ile
365 370 375
Val Phe Pro Val Leu Phe Gly Thr Phe Asn Leu Val Tyr Trp Ala
380 385 390
Thr Tyr Leu Asn Arg Glu Pro Val Leu Gly Val Ser Pro
395 400

<210> 43
<211> 66
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506960CD1

<400> 43
Met Lys Thr Lys Leu Asn Ile Tyr Asn Met Gln Phe Leu Leu Phe
1 5 10 15
Val Phe Leu Val Trp Asp Pro Ala Arg Leu Val Leu Ala Asn Ile
20 25 30
Gln Glu Asp Glu Ala Lys Asn Asn Ile Thr Ile Phe Thr Arg Ile
35 40 45
Leu Asp Arg Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly
50 55 60
Leu Gly Gly Ile Tyr Asn
65

<210> 44
<211> 89

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510540CD1

<400> 44
Met Thr Glu Asp Lys Val Thr Gly Thr Leu Val Phe Thr Val Ile
1 5 10 15
Thr Ala Val Leu Gly Ser Phe Gln Phe Gly Tyr Asp Ile Gly Val
20 25 30
Ile Asn Ala Pro Gln Gln Asn Gln Ser His Val Ser Ser Lys His
35 40 45
Ser Val Ile Ser Trp Ser Ser Leu Asp Gly Val Phe Lys Ile Gly
50 55 60
Thr Ile Ser Tyr Thr Tyr Asn Cys Trp Lys Lys His Ile Arg Thr
65 70 75
Ile Leu Trp Ala Asn Phe Arg Pro Gly Ser Tyr Val Tyr Arg
80 85

<210> 45
<211> 146
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510545CD1

<400> 45
Met Glu Asn Ala His Thr Lys Thr Val Glu Glu Val Leu Gly His
1 5 10 15
Phe Gly Val Asn Glu Ser Glu Ser Val Ser Val Ile Lys His Thr
20 25 30
Asp Pro Val Pro Asp Pro Arg Ala Val Asn Gln Asp Lys Lys Asn
35 40 45
Met Leu Phe Ser Val Ala Leu Ala Val Ala Ala Ile Pro Glu Gly
50 55 60
Leu Pro Ala Val Ile Thr Thr Cys Leu Ala Leu Gly Thr Arg Arg
65 70 75
Met Ala Lys Lys Asn Ala Ile Val Arg Ser Leu Pro Ser Val Glu
80 85 90
Thr Leu Gly Cys Thr Ser Val Ile Cys Ser Asp Lys Thr Gly Thr
95 100 105
Leu Thr Thr Asn Gln Met Ser Val Cys Arg Met Phe Ile Leu Asp
110 115 120
Arg Val Glu Asp His Thr Ala Glu Arg Asp Pro Val Ala Asp Gly
125 130 135
Ala Glu Asn Leu Leu Ala Arg Asp Ser His Gly
140 145

<210> 46
<211> 353
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510654CD1

<400> 46
Met Thr Pro Glu Asp Pro Glu Glu Thr Gln Pro Leu Leu Gly Pro
1 5 10 15
Pro Gly Gly Ser Ala Pro Arg Gly Arg Arg Val Phe Leu Ala Ala
20 25 30
Phe Ala Ala Ala Leu Gly Pro Leu Ser Phe Gly Phe Ala Leu Gly
35 40 45
Tyr Ser Ser Pro Ala Ile Pro Ser Leu Gln Arg Ala Ala Pro Pro
50 55 60
Ala Pro Arg Leu Asp Asp Ala Ala Ala Ser Trp Phe Gly Ala Val
65 70 75
Val Thr Leu Gly Ala Ala Ala Gly Gly Val Leu Gly Gly Trp Leu
80 85 90
Val Asp Arg Ala Gly Arg Lys Leu Ser Leu Leu Leu Cys Ser Val
95 100 105
Pro Phe Val Ala Gly Phe Ala Val Ile Thr Ala Ala Gln Asp Val
110 115 120
Trp Met Leu Leu Gly Gly Arg Leu Leu Thr Gly Leu Ala Cys Gly
125 130 135
Val Ala Ser Leu Val Ala Pro Val Tyr Ile Ser Glu Ile Ala Tyr
140 145 150
Pro Ala Val Arg Gly Leu Leu Gly Ser Cys Val Gln Leu Met Val
155 160 165
Val Val Gly Ile Leu Leu Ala Tyr Leu Ala Gly Trp Val Leu Glu
170 175 180
Trp Arg Trp Leu Ala Val Leu Gly Cys Val Pro Pro Ser Leu Met
185 190 195
Leu Leu Leu Met Cys Phe Met Pro Glu Thr Pro Arg Phe Leu Leu
200 205 210
Thr Gln His Arg Arg Gln Glu Ala Met Ala Ala Leu Arg Phe Leu
215 220 225
Trp Gly Ser Glu Gln Gly Trp Glu Asp Pro Pro Ile Gly Ala Glu
230 235 240
Gln Ser Phe His Leu Ala Leu Leu Arg Gln Pro Gly Ile Tyr Lys
245 250 255
Pro Phe Ile Ile Gly Val Ser Leu Met Ala Phe Gln Gln Leu Ser
260 265 270
Gly Val Asn Ala Val Met Phe Tyr Ala Glu Thr Ile Phe Glu Glu
275 280 285
Ala Lys Phe Lys Asp Ser Ser Leu Ala Ser Val Val Val Gly Val
290 295 300
Ile Gln Val Leu Phe Thr Ala Val Ala Ala Leu Ile Met Asp Arg
305 310 315
Ala Gly Arg Arg Leu Leu Leu Val Leu Ser Gly Gly Pro Gln Ala
320 325 330
Leu Trp Ser Leu Leu Ala Cys Leu Arg Phe Leu His Leu Gln Cys
335 340 345
Pro Phe His Phe Val Leu Cys Pro
350

<210> 47
<211> 1155

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510660CD1

<400> 47
Met Ala Ala Ala Ala Ala Val Gly Asn Ala Val Pro Cys Gly Ala
1 5 10 15
Arg Pro Cys Gly Val Arg Pro Asp Gly Gln Pro Lys Pro Gly Pro
20 25 30
Gln Pro Arg Ala Leu Leu Ala Ala Gly Pro Ala Leu Ile Ala Asn
35 40 45
Gly Asp Glu Leu Val Ala Ala Val Trp Pro Tyr Arg Arg Leu Ala
50 55 60
Leu Leu Arg Arg Leu Thr Val Leu Pro Phe Ala Gly Leu Leu Tyr
65 70 75
Pro Ala Trp Leu Gly Ala Ala Ala Ala Gly Cys Trp Gly Trp Gly
80 85 90
Ser Ser Trp Val Gln Ile Pro Glu Ala Ala Leu Leu Val Leu Ala
95 100 105
Thr Ile Cys Leu Ala His Ala Leu Thr Val Leu Ser Gly His Trp
110 115 120
Ser Val His Ala His Cys Ala Leu Thr Cys Thr Pro Glu Tyr Asp
125 130 135
Pro Ser Lys Ala Thr Phe Val Lys Val Val Pro Thr Pro Asn Asn
140 145 150
Gly Ser Thr Glu Leu Val Ala Leu His Arg Asn Glu Gly Glu Asp
155 160 165
Gly Leu Glu Val Leu Ser Phe Glu Phe Gln Lys Ile Lys Tyr Ser
170 175 180
Tyr Asp Ala Leu Glu Lys Lys Gln Phe Leu Pro Val Ala Phe Pro
185 190 195
Val Gly Asn Ala Phe Ser Tyr Tyr Gln Ser Asn Arg Gly Phe Gln
200 205 210
Glu Asp Ser Glu Ile Arg Ala Ala Glu Lys Lys Phe Gly Ser Asn
215 220 225
Lys Ala Glu Met Val Val Pro Asp Phe Ser Glu Leu Phe Lys Glu
230 235 240
Arg Ala Thr Ala Pro Phe Phe Val Phe Gln Val Phe Cys Val Gly
245 250 255
Leu Trp Cys Leu Asp Glu Tyr Trp Tyr Tyr Ser Val Phe Thr Leu
260 265 270
Ser Met Leu Val Ala Phe Glu Ala Ser Leu Val Gln Gln Gln Met
275 280 285
Arg Asn Met Ser Glu Ile Arg Lys Met Gly Asn Lys Pro His Met
290 295 300
Ile Gln Val Tyr Arg Ser Arg Lys Trp Arg Pro Ile Ala Ser Asp
305 310 315
Glu Ile Val Pro Gly Asp Ile Val Ser Ile Gly Arg Ser Pro Gln
320 325 330
Glu Asn Leu Val Pro Cys Asp Val Leu Leu Leu Arg Gly Arg Cys
335 340 345
Ile Val Asp Glu Ala Met Leu Thr Gly Glu Ser Val Pro Gln Met
350 355 360

Lys Glu Pro Ile Glu Asp Leu Ser Pro Asp Arg Val Leu Asp Leu
365 370 375
Gln Ala Asp Ser Arg Leu His Val Ile Phe Gly Gly Thr Lys Val
380 385 390
Val Gln His Ile Pro Pro Gln Lys Ala Thr Thr Gly Leu Lys Pro
395 400 405
Val Asp Ser Gly Cys Val Ala Tyr Val Leu Arg Thr Gly Phe Asn
410 415 420
Thr Ser Gln Gly Lys Leu Leu Arg Thr Ile Leu Phe Gly Val Lys
425 430 435
Arg Val Thr Ala Asn Asn Leu Glu Thr Phe Ile Phe Ile Leu Phe
440 445 450
Leu Leu Val Phe Ala Ile Ala Ala Ala Ala Tyr Val Trp Ile Glu
455 460 465
Gly Thr Lys Asp Pro Ser Arg Asn Arg Tyr Lys Leu Phe Leu Glu
470 475 480
Cys Thr Leu Ile Leu Thr Ser Val Val Pro Pro Glu Leu Pro Ile
485 490 495
Glu Leu Ser Leu Ala Val Asn Thr Ser Leu Ile Ala Leu Ala Lys
500 505 510
Leu Tyr Met Tyr Cys Thr Glu Pro Phe Arg Ile Pro Phe Ala Gly
515 520 525
Lys Val Glu Val Cys Cys Phe Asp Lys Thr Gly Thr Leu Thr Ser
530 535 540
Asp Ser Leu Val Val Arg Gly Val Ala Gly Leu Arg Asp Gly Lys
545 550 555
Glu Val Thr Pro Val Ser Ser Ile Pro Val Glu Thr His Arg Ala
560 565 570
Leu Ala Ser Cys His Ser Leu Met Gln Leu Asp Asp Gly Thr Leu
575 580 585
Val Gly Asp Pro Leu Glu Lys Ala Met Leu Thr Ala Val Asp Trp
590 595 600
Thr Leu Thr Lys Asp Glu Lys Val Phe Pro Arg Ser Ile Lys Thr
605 610 615
Gln Gly Leu Lys Ile His Gln Arg Phe His Phe Ala Ser Ala Leu
620 625 630
Lys Arg Met Ser Val Leu Ala Ser Tyr Glu Lys Leu Gly Ser Thr
635 640 645
Asp Leu Cys Tyr Ile Ala Ala Val Lys Gly Ala Pro Glu Thr Leu
650 655 660
His Ser Met Phe Ser Gln Cys Pro Pro Asp Tyr His His Ile His
665 670 675
Thr Glu Ile Ser Arg Glu Gly Ala Arg Val Leu Ala Leu Gly Tyr
680 685 690
Lys Glu Leu Gly His Leu Thr His Gln Gln Val Val Met Ile Thr
695 700 705
Gly Asp Asn Pro Leu Thr Ala Cys His Val Ala Gln Glu Leu His
710 715 720
Phe Ile Glu Lys Ala His Thr Leu Ile Leu Gln Pro Pro Ser Glu
725 730 735
Lys Gly Arg Gln Cys Glu Trp Arg Ser Ile Asp Gly Ser Ile Val
740 745 750
Leu Pro Leu Ala Arg Gly Ser Pro Lys Ala Leu Ala Leu Glu Tyr
755 760 765
Ala Leu Cys Leu Thr Gly Asp Gly Leu Ala His Leu Gln Ala Thr
770 775 780

Asp Pro Gln Gln Leu Leu Arg Leu Ile Pro His Val Gln Val Phe
785 790 795
Ala Arg Val Ala Pro Lys Gln Lys Glu Phe Val Ile Thr Ser Leu
800 805 810
Lys Glu Leu Gly Tyr Val Thr Leu Met Cys Gly Asp Gly Thr Asn
815 820 825
Asp Val Gly Ala Leu Lys His Ala Asp Val Gly Val Ala Leu Leu
830 835 840
Ala Asn Ala Pro Glu Arg Val Val Glu Arg Arg Arg Arg Pro Arg
845 850 855
Asp Ser Pro Thr Leu Ser Asn Ser Gly Ile Arg Ala Thr Ser Arg
860 865 870
Thr Ala Lys Gln Arg Ser Gly Leu Pro Pro Ser Glu Glu Gln Pro
875 880 885
Thr Ser Gln Arg Asp Arg Leu Ser Gln Val Leu Arg Asp Leu Glu
890 895 900
Asp Glu Ser Thr Pro Ile Val Lys Leu Gly Asp Ala Ser Ile Ala
905 910 915
Ala Pro Phe Thr Ser Lys Leu Ser Ser Ile Gln Cys Ile Cys His
920 925 930
Val Ile Lys Gln Gly Arg Cys Thr Leu Val Thr Thr Leu Gln Met
935 940 945
Phe Lys Ile Leu Ala Leu Asn Ala Leu Ile Leu Ala Tyr Ser Gln
950 955 960
Ser Val Leu Tyr Leu Glu Gly Val Lys Phe Ser Asp Phe Gln Ala
965 970 975
Thr Leu Gln Gly Leu Leu Leu Ala Gly Cys Phe Leu Phe Ile Ser
980 985 990
Arg Ser Lys Pro Leu Lys Thr Leu Ser Arg Glu Arg Pro Leu Pro
995 1000 1005
Asn Ile Phe Asn Leu Tyr Thr Ile Leu Thr Val Met Leu Gln Phe
1010 1015 1020
Phe Val His Phe Leu Ser Leu Val Tyr Leu Tyr Arg Glu Ala Gln
1025 1030 1035
Ala Arg Ser Pro Glu Lys Gln Glu Gln Phe Val Asp Leu Tyr Lys
1040 1045 1050
Glu Phe Glu Pro Ser Leu Val Asn Ser Thr Val Tyr Ile Met Ala
1055 1060 1065
Met Ala Met Gln Met Ala Thr Phe Ala Ile Asn Tyr Lys Gly Pro
1070 1075 1080
Pro Phe Met Glu Ser Leu Pro Glu Asn Lys Pro Leu Val Trp Ser
1085 1090 1095
Leu Ala Val Ser Leu Leu Ala Ile Ile Gly Leu Leu Leu Gly Ser
1100 1105 1110
Ser Pro Asp Phe Asn Ser Gln Phe Gly Leu Val Asp Ile Pro Val
1115 1120 1125
Glu Val Leu Leu Leu Asp Phe Cys Leu Ala Leu Leu Ala Asp Arg
1130 1135 1140
Val Leu Gln Phe Phe Leu Gly Thr Pro Lys Leu Lys Val Pro Ser
1145 1150 1155

<210> 48
<211> 606
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510661CD1

<400> 48
Met Ala Ala Ala Ala Ala Val Gly Asn Ala Val Pro Cys Gly Ala
1               5                   10                  15
Arg Pro Cys Gly Val Arg Pro Asp Gly Gln Pro Lys Pro Gly Pro
                20                  25                  30
Gln Pro Arg Ala Leu Leu Ala Ala Gly Pro Ala Leu Ile Ala Asn
                35                  40                  45
Gly Asp Glu Leu Val Ala Ala Val Trp Pro Tyr Arg Arg Leu Ala
                50                  55                  60
Leu Leu Arg Arg Leu Thr Val Leu Pro Phe Ala Gly Leu Leu Tyr
                65                  70                  75
Pro Ala Trp Leu Gly Ala Ala Ala Ala Gly Cys Trp Gly Trp Gly
                80                  85                  90
Ser Ser Trp Val Gln Ile Pro Glu Ala Ala Leu Leu Val Leu Ala
                95                  100                 105
Thr Ile Cys Leu Ala His Ala Leu Thr Val Leu Ser Gly His Trp
                110                 115                 120
Ser Val His Ala His Cys Ala Leu Thr Cys Thr Pro Glu Tyr Asp
                125                 130                 135
Pro Ser Lys Ala Thr Phe Val Lys Val Val Pro Thr Pro Asn Asn
                140                 145                 150
Gly Ser Thr Glu Leu Val Ala Leu His Arg Asn Glu Gly Glu Asp
                155                 160                 165
Gly Leu Glu Val Leu Ser Phe Glu Phe Gln Lys Ile Lys Tyr Ser
                170                 175                 180
Tyr Asp Ala Leu Glu Lys Lys Gln Phe Leu Pro Val Ala Phe Pro
                185                 190                 195
Val Gly Asn Ala Phe Ser Tyr Tyr Gln Ser Asn Arg Gly Phe Gln
                200                 205                 210
Glu Asp Ser Glu Ile Arg Ala Ala Glu Lys Lys Phe Gly Ser Asn
                215                 220                 225
Lys Ala Glu Met Val Val Pro Asp Phe Ser Glu Leu Phe Lys Glu
                230                 235                 240
Arg Ala Thr Ala Pro Phe Phe Val Phe Gln Val Phe Cys Val Gly
                245                 250                 255
Leu Trp Cys Leu Asp Glu Tyr Trp Tyr Tyr Ser Val Phe Thr Leu
                260                 265                 270
Ser Met Leu Val Ala Phe Glu Ala Ser Leu Val Gln Gln Gln Met
                275                 280                 285
Arg Asn Met Ser Glu Ile Arg Lys Met Gly Asn Lys Pro His Met
                290                 295                 300
Ile Gln Val Tyr Arg Ser Arg Lys Trp Arg Pro Ile Ala Ser Asp
                305                 310                 315
Glu Ile Val Pro Gly Asp Ile Val Ser Ile Gly Arg Ser Pro Gln
                320                 325                 330
Glu Asn Leu Val Pro Cys Asp Val Leu Leu Leu Arg Gly Arg Cys
                335                 340                 345
Ile Val Asp Glu Ala Met Leu Thr Gly Glu Ser Val Pro Gln Met
                350                 355                 360
Lys Glu Pro Ile Glu Asp Leu Ser Pro Asp Arg Val Leu Asp Leu
                365                 370                 375
Gln Ala Asp Ser Arg Leu His Val Ile Phe Gly Gly Thr Lys Val

```
                380                 385                 390
Val Gln His Ile Pro Pro Gln Lys Ala Thr Thr Gly Leu Lys Pro
                395                 400                 405
Val Asp Ser Gly Cys Val Ala Tyr Val Leu Arg Thr Gly Phe Asn
                410                 415                 420
Thr Ser Gln Gly Lys Leu Leu Arg Thr Ile Leu Phe Gly Val Lys
                425                 430                 435
Arg Val Thr Ala Asn Asn Leu Glu Thr Phe Ile Phe Ile Leu Phe
                440                 445                 450
Leu Leu Val Phe Ala Ile Ala Ala Ala Ala Tyr Val Trp Ile Glu
                455                 460                 465
Gly Thr Lys Asp Pro Ser Arg Asn Arg Tyr Lys Leu Phe Leu Glu
                470                 475                 480
Cys Thr Leu Ile Leu Thr Ser Val Val Pro Pro Glu Leu Pro Ile
                485                 490                 495
Glu Leu Ser Leu Ala Val Asn Thr Ser Leu Ile Ala Leu Ala Lys
                500                 505                 510
Leu Tyr Met Tyr Cys Thr Glu Pro Phe Arg Ile Pro Phe Ala Gly
                515                 520                 525
Lys Val Glu Val Cys Cys Phe Asp Lys Thr Gly Thr Leu Thr Ser
                530                 535                 540
Asp Ser Leu Val Val Arg Gly Val Ala Gly Leu Arg Asp Gly Lys
                545                 550                 555
Glu Val Thr Pro Val Ser Ser Ile Pro Val Glu Thr His Arg Ala
                560                 565                 570
Leu Ala Ser Cys His Ser Leu Met Gln Leu Asp Asp Gly Thr Leu
                575                 580                 585
Val Gly Asp Pro Leu Glu Lys Ala Met Leu Thr Ala Val Asp Trp
                590                 595                 600
Thr Leu Thr Lys Val Pro
                605
```

<210> 49
<211> 462
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510680CD1

<400> 49

```
Met Ala Thr Lys Pro Thr Glu Pro Val Thr Ile Leu Ser Leu Arg
  1               5                  10                  15
Lys Leu Ser Leu Gly Thr Ala Glu Pro Gln Val Lys Glu Pro Lys
                 20                  25                  30
Thr Phe Thr Val Glu Asp Ala Val Glu Thr Ile Gly Phe Gly Arg
                 35                  40                  45
Phe His Ile Ala Leu Phe Leu Ile Met Gly Ser Thr Gly Val Val
                 50                  55                  60
Glu Ala Met Glu Ile Met Leu Ile Ala Val Val Ser Pro Val Ile
                 65                  70                  75
Arg Cys Glu Trp Gln Leu Glu Asn Trp Gln Val Ala Leu Val Thr
                 80                  85                  90
Thr Met Val Phe Phe Gly Tyr Met Val Phe Ser Ile Leu Phe Gly
                 95                 100                 105
```

```
Leu Leu Ala Asp Arg Tyr Gly Arg Trp Lys Ile Leu Leu Ile Ser
                110                 115                 120
Phe Leu Trp Gly Ala Tyr Phe Ser Leu Leu Thr Ser Phe Ala Pro
                125                 130                 135
Ser Tyr Ile Trp Phe Val Phe Leu Arg Thr Met Val Gly Cys Gly
                140                 145                 150
Val Ser Gly His Ser Gln Gly Leu Ile Ile Lys Thr Glu Phe Leu
                155                 160                 165
Pro Thr Lys Tyr Arg Gly Tyr Met Leu Pro Leu Ser Gln Val Phe
                170                 175                 180
Trp Leu Ala Gly Ser Leu Leu Ile Ile Gly Leu Ala Ser Val Ile
                185                 190                 195
Ile Pro Thr Ile Gly Trp Arg Trp Leu Ile Arg Val Ala Ser Ile
                200                 205                 210
Pro Gly Ile Ile Leu Ile Val Ala Phe Lys Phe Ile Pro Glu Ser
                215                 220                 225
Ala Arg Phe Asn Val Ser Thr Gly Asn Thr Arg Ala Ala Leu Ala
                230                 235                 240
Thr Leu Glu Arg Val Ala Lys Met Asn Arg Ser Val Met Pro Glu
                245                 250                 255
Gly Lys Leu Val Glu Pro Val Leu Glu Lys Arg Gly Arg Phe Ala
                260                 265                 270
Asp Leu Leu Asp Ala Lys Tyr Leu Arg Thr Thr Leu Gln Ile Trp
                275                 280                 285
Val Ile Trp Leu Gly Ile Ser Phe Ala Tyr Tyr Gly Val Ile Leu
                290                 295                 300
Ala Ser Ala Glu Leu Leu Glu Arg Asp Leu Val Cys Gly Ser Lys
                305                 310                 315
Ser Asp Ser Ala Val Val Val Thr Gly Gly Asp Ser Gly Glu Ser
                320                 325                 330
Gln Ser Pro Cys Tyr Cys His Met Phe Ala Pro Ser Asp Tyr Arg
                335                 340                 345
Thr Met Ile Ile Ser Thr Ile Gly Glu Ile Ala Leu Asn Pro Leu
                350                 355                 360
Asn Ile Leu Gly Ile Asn Phe Leu Gly Arg Arg Leu Ser Leu Ser
                365                 370                 375
Ile Thr Met Gly Cys Thr Ala Leu Phe Phe Leu Leu Leu Asn Ile
                380                 385                 390
Cys Thr Ser Ser Ala Gly Leu Ile Gly Phe Leu Phe Met Leu Arg
                395                 400                 405
Ala Leu Val Ala Ala Asn Phe Asn Thr Val Tyr Ile Tyr Thr Ala
                410                 415                 420
Glu Val Leu Met Ser Ala Ser Ile Leu Gly Ala Leu Cys Leu Phe
                425                 430                 435
Ser Ser Val Cys Val Val Cys Ala Ile Ser Ala Phe Thr Leu Pro
                440                 445                 450
Ile Glu Thr Lys Gly Arg Ala Leu Gln Gln Ile Lys
                455                 460
```

<210> 50
<211> 366
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7505145CD1

<400> 50
Met Gly Trp Gly Gly Gly Gly Gly Cys Thr Pro Arg Pro Pro Ile
1 5 10 15
His Gln Gln Pro Pro Glu Arg Arg Val Val Thr Val Val Phe Leu
20 25 30
Gly Leu Leu Leu Asp Leu Leu Ala Phe Thr Leu Leu Leu Pro Leu
35 40 45
Leu Pro Gly Leu Leu Glu Ser His Gly Arg Ala His Asp Pro Leu
50 55 60
Tyr Gly Ser Trp Gln Gly Gly Val Asp Trp Phe Ala Thr Ala Ile
65 70 75
Gly Met Pro Val Glu Lys Arg Tyr Asn Ser Val Leu Phe Gly Gly
80 85 90
Leu Ile Gly Ser Ala Phe Ser Val Leu Gln Phe Leu Cys Ala Pro
95 100 105
Leu Thr Gly Ala Thr Ser Asp Cys Leu Gly Arg Arg Pro Val Met
110 115 120
Leu Leu Cys Leu Met Gly Val Ala Thr Ser Tyr Ala Val Trp Ala
125 130 135
Thr Ser Arg Ser Phe Ala Ala Phe Leu Ala Ser Arg Leu Ile Gly
140 145 150
Gly Ile Ser Lys Gly Asn Val Ser Leu Ser Thr Ala Ile Val Ala
155 160 165
Asp Leu Gly Ser Pro Leu Ala Arg Ser Gln Gly Met Ala Val Ile
170 175 180
Gly Val Ala Phe Ser Leu Gly Phe Thr Leu Gly Pro Met Leu Gly
185 190 195
Ala Ser Leu Pro Leu Glu Met Ala Pro Trp Phe Ala Leu Leu Phe
200 205 210
Ala Ala Ser Asp Leu Leu Phe Ile Phe Cys Phe Leu Pro Glu Thr
215 220 225
Leu Pro Leu Glu Lys Arg Ala Pro Ser Ile Ala Leu Gly Phe Arg
230 235 240
Asp Ala Ala Asp Leu Leu Ser Pro Leu Ala Leu Leu Arg Phe Ser
245 250 255
Ala Val Ala Arg Gly Gln Asp Pro Pro Ser Gly Asp Arg Leu Ser
260 265 270
Ser Leu Arg Arg Leu Gly Leu Val Tyr Phe Leu Tyr Leu Phe Leu
275 280 285
Phe Ser Gly Leu Glu Tyr Thr Leu Ser Phe Leu Thr His Gln Arg
290 295 300
Phe Gln Phe Ser Arg Pro Ser Cys Cys Trp Cys Pro Pro Ser Ser
305 310 315
Ser Ser Ala Gly Asp Val Leu Cys Pro Cys Trp Ala Trp Gly Cys
320 325 330
Cys Ser Thr Pro Leu Pro Pro Pro Leu Trp Cys Pro Ala Cys Pro
335 340 345
Pro Trp Ser Leu Ala Met Ala His Gln Gly Arg Arg Ala Arg Ser
350 355 360
Trp Val His Cys Ala Ala
365

<210> 51
<211> 295

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505162CD1

<400> 51
```
Met Ala Ala Gln Gly Tyr Gly Tyr Tyr Arg Thr Val Ile Phe Ser
  1               5                  10                  15
Ala Met Phe Gly Gly Tyr Ser Leu Tyr Tyr Phe Asn Arg Lys Thr
                 20                  25                  30
Phe Ser Phe Val Met Pro Ser Leu Val Glu Glu Ile Pro Leu Asp
                 35                  40                  45
Lys Asp Asp Leu Gly Phe Ile Thr Ser Ser Gln Ser Ala Ala Tyr
                 50                  55                  60
Ala Ile Ser Lys Phe Val Ser Gly Val Leu Ser Asp Gln Met Ser
                 65                  70                  75
Ala Arg Trp Leu Phe Ser Ser Gly Leu Leu Leu Val Gly Leu Val
                 80                  85                  90
Asn Ile Phe Phe Ala Trp Ser Ser Thr Val Pro Val Phe Ala Ala
                 95                 100                 105
Leu Trp Phe Leu Asn Gly Leu Ala Gln Gly Leu Gly Trp Pro Pro
                110                 115                 120
Cys Gly Lys Val Leu Arg Lys Trp Phe Glu Pro Ser Gln Phe Gly
                125                 130                 135
Thr Trp Trp Ala Ile Leu Ser Thr Ser Met Asn Leu Ala Gly Gly
                140                 145                 150
Leu Gly Pro Ile Leu Ala Thr Ile Leu Ala Gln Ser Tyr Ser Trp
                155                 160                 165
Arg Ser Thr Leu Ala Leu Ser Gly Ala Leu Cys Val Val Val Ser
                170                 175                 180
Phe Leu Cys Leu Leu Leu Ile His Asn Glu Pro Ala Asp Val Gly
                185                 190                 195
Leu Arg Asn Leu Asp Pro Met Pro Ser Glu Gly Lys Lys Gly Ser
                200                 205                 210
Leu Lys Glu Glu Ser Thr Leu Gln Glu Leu Leu Leu Ser Pro Tyr
                215                 220                 225
Leu Trp Val Leu Ser Thr Gly Tyr Leu Val Val Phe Gly Val Lys
                230                 235                 240
Thr Cys Cys Thr Asp Trp Gly Gln Phe Phe Leu Ile Gln Glu Lys
                245                 250                 255
Gly Gln Ser Ala Leu Val Gly Gly Thr Val Gln Leu Arg Glu Pro
                260                 265                 270
Ser Pro Trp Pro Val Ala Val His Asp Gly Trp His Asp Ser Val
                275                 280                 285
His Val Pro Leu Pro Gly Asn Ser Asp Gln
                290                 295
```

<210> 52
<211> 229
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7505469CD1

<400> 52
Met Glu Ala Arg Glu Pro Gly Arg Pro Thr Pro Thr Tyr His Leu
1 5 10 15
Val Pro Asn Thr Ser Gln Ser Gln Val Glu Glu Asp Val Ser Ser
20 25 30
Pro Pro Gln Arg Ser Ser Glu Thr Met Gln Leu Lys Lys Glu Ile
35 40 45
Ser Leu Leu Asn Gly Val Ser Leu Val Val Gly Asn Met Ile Gly
50 55 60
Ser Gly Ile Phe Val Ser Pro Lys Gly Val Leu Val His Thr Ala
65 70 75
Ser Tyr Gly Met Ser Leu Ile Val Trp Ala Ile Gly Gly Leu Phe
80 85 90
Ser Val Val Gly Ala Leu Cys Tyr Ala Glu Leu Gly Thr Thr Ile
95 100 105
Thr Lys Ser Gly Ala Ser Tyr Ala Tyr Ile Leu Glu Ala Phe Gly
110 115 120
Gly Phe Ile Ala Phe Ile Arg Leu Trp Val Ser Leu Leu Val Val
125 130 135
Glu Pro Thr Gly Gln Ala Ile Ile Ala Ile Thr Phe Ala Asn Tyr
140 145 150
Ile Ile Gln Pro Ser Phe Pro Ser Cys Asp Pro Pro Tyr Leu Ala
155 160 165
Cys Arg Leu Leu Ala Ala Ala Cys Ile Cys Leu Leu Thr Phe Val
170 175 180
Asn Cys Ala Tyr Val Lys Trp Gly Thr Arg Val Gln Asp Thr Phe
185 190 195
Thr Tyr Ala Lys Val Val Ala Leu Ile Ala Ile Ile Val Met Gly
200 205 210
Leu Val Lys Leu Cys Gln Glu Ile Cys Pro Trp Pro Leu Gly Phe
215 220 225
Leu Cys Gln Leu

<210> 53
<211> 637
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505475CD1

<400> 53
Met Asn Met Lys Gln Lys Ser Val Tyr Gln Gln Thr Lys Ala Leu
1 5 10 15
Leu Cys Lys Asn Phe Leu Lys Lys Trp Arg Met Lys Arg Glu Ser
20 25 30
Leu Leu Glu Trp Gly Leu Ser Ile Leu Leu Gly Leu Cys Ile Ala
35 40 45
Leu Phe Ser Ser Ser Met Arg Asn Val Gln Phe Pro Gly Met Ala
50 55 60
Pro Gln Asn Leu Gly Arg Val Asp Lys Phe Asn Ser Ser Ser Leu
65 70 75

Met Val Val Tyr Thr Pro Ile Ser Asn Leu Thr Gln Gln Ile Met
80 85 90
Asn Lys Thr Ala Leu Ala Pro Leu Leu Lys Gly Thr Ser Val Ile
95 100 105
Gly Ala Pro Asn Lys Thr His Met Asp Glu Ile Leu Leu Glu Asn
110 115 120
Leu Pro Tyr Ala Met Gly Ile Ile Phe Asn Glu Thr Phe Ser Tyr
125 130 135
Lys Leu Ile Phe Phe Gln Gly Tyr Asn Ser Pro Leu Trp Lys Glu
140 145 150
Asp Phe Ser Ala His Cys Trp Asp Gly Tyr Gly Glu Phe Ser Cys
155 160 165
Thr Leu Thr Lys Tyr Trp Asn Arg Gly Phe Val Ala Leu Gln Thr
170 175 180
Ala Ile Asn Thr Ala Ile Ile Glu Ile Thr Thr Asn His Pro Val
185 190 195
Met Glu Glu Leu Met Ser Val Thr Ala Ile Thr Met Lys Thr Leu
200 205 210
Pro Phe Ile Thr Lys Asn Leu Leu His Asn Glu Met Phe Ile Leu
215 220 225
Phe Phe Leu Leu His Phe Ser Pro Leu Val Tyr Phe Ile Ser Leu
230 235 240
Asn Val Thr Lys Glu Arg Lys Lys Ser Lys Asn Leu Met Lys Met
245 250 255
Met Gly Leu Gln Asp Ser Ala Phe Trp Leu Ser Trp Gly Leu Ile
260 265 270
Tyr Ala Gly Phe Ile Phe Ile Ile Ser Ile Phe Ile Thr Ile Ile
275 280 285
Ile Thr Phe Thr Gln Ile Ile Val Met Thr Gly Phe Met Val Ile
290 295 300
Phe Ile Pro Phe Phe Leu Tyr Gly Leu Ser Leu Val Ala Leu Val
305 310 315
Phe Leu Leu Ser Val Leu Leu Lys Lys Ala Val Leu Thr Asn Leu
320 325 330
Val Val Phe Leu Leu Thr Leu Phe Trp Gly Cys Leu Gly Phe Thr
335 340 345
Val Phe Tyr Glu Gln Leu Pro Ser Ser Leu Glu Trp Ile Leu Asn
350 355 360
Ile Cys Ser Pro Phe Ala Phe Thr Thr Gly Met Ile Gln Ile Ile
365 370 375
Lys Leu Asp Tyr Asn Leu Asn Gly Val Ile Phe Pro Asp Pro Ser
380 385 390
Gly Asp Ser Tyr Thr Met Ile Ala Thr Phe Ser Met Leu Leu Leu
395 400 405
Asp Gly Leu Ile Tyr Leu Leu Leu Ala Leu Tyr Phe Asp Lys Ile
410 415 420
Leu Pro Tyr Gly Asp Glu Arg His Tyr Ser Pro Leu Phe Phe Leu
425 430 435
Asn Ser Ser Ser Cys Phe Gln His Gln Arg Thr Asn Ala Lys Val
440 445 450
Ile Glu Lys Glu Ile Asp Ala Glu His Pro Ser Asp Asp Tyr Phe
455 460 465
Glu Pro Val Ala Pro Glu Phe Gln Gly Lys Glu Ala Ile Arg Ile
470 475 480
Arg Asn Val Lys Lys Glu Tyr Lys Gly Lys Ser Gly Lys Val Glu
485 490 495

Ala Leu Lys Gly Leu Leu Phe Asp Ile Tyr Glu Gly Gln Ile Thr
500 505 510
Ala Ile Leu Gly His Ser Gly Ala Gly Lys Ser Ser Leu Leu Asn
515 520 525
Ile Leu Asn Gly Leu Ser Val Pro Thr Glu Gly Ser Val Thr Ile
530 535 540
Tyr Asn Lys Asn Leu Ser Glu Met Gln Asp Leu Glu Glu Ile Arg
545 550 555
Lys Ile Thr Gly Val Cys Pro Gln Phe Asn Val Gln Phe Asp Ile
560 565 570
Leu Thr Val Lys Glu Asn Leu Ser Leu Phe Ala Lys Ile Lys Gly
575 580 585
Ile His Leu Lys Glu Val Glu Gln Glu Val Gln Arg Ile Leu Leu
590 595 600
Glu Leu Asp Met Gln Asn Ile Gln Asp Asn Leu Ala Lys His Leu
605 610 615
Ser Glu Gly Gln Lys Arg Lys Leu Thr Phe Gly Ile Thr Ile Leu
620 625 630
Gly Asp Pro Gln Ile Glu Lys
635

<210> 54
<211> 90
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505568CD1

<400> 54
Met Ala Arg Lys Gln Asn Arg Asn Ser Lys Glu Leu Gly Leu Val
1 5 10 15
Pro Leu Thr Asp Asp Thr Ser His Ala Gly Pro Pro Gly Pro Gly
20 25 30
Arg Ala Leu Leu Glu Cys Asp His Leu Arg Ser Gly Val Pro Gly
35 40 45
Gly Arg Arg Arg Lys Asp Trp Ser Cys Ser Leu Leu Val Ala Ser
50 55 60
Leu Ala Gly Ala Phe Gly Ser Ser Phe Leu Tyr Gly Tyr Asn Leu
65 70 75
Ser Val Val Asn Ala Pro Thr Pro Glu Ala His Phe Ala Gly Gln
80 85 90

<210> 55
<211> 327
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506953CD1

<400> 55
Met Lys Thr Lys Leu Asn Ile Tyr Asn Met Gln Phe Leu Leu Phe

```
1               5                   10                  15
Val Phe Leu Val Trp Asp Pro Ala Arg Leu Val Leu Ala Asn Ile
                20                  25                  30
Gln Glu Asp Glu Ala Lys Asn Asn Ile Thr Ile Phe Thr Arg Ile
                35                  40                  45
Leu Asp Arg Leu Leu Asp Gly Tyr Asp Asn Arg Leu Arg Pro Gly
                50                  55                  60
Leu Gly Asp Ala Tyr Thr Thr Ser Glu Val Thr Tyr Ile Trp Thr
                65                  70                  75
Tyr Asn Ala Ser Asp Ser Val Gln Val Ala Pro Asp Gly Ser Arg
                80                  85                  90
Leu Asn Gln Tyr Asp Leu Leu Gly Gln Ser Ile Gly Lys Glu Thr
                95                  100                 105
Ile Lys Ser Ser Thr Gly Glu Tyr Thr Val Met Thr Ala His Phe
                110                 115                 120
His Leu Lys Arg Lys Ile Gly Tyr Phe Val Ile Gln Thr Tyr Leu
                125                 130                 135
Pro Cys Ile Met Thr Val Ile Leu Ser Gln Val Ser Phe Trp Leu
                140                 145                 150
Asn Arg Glu Ser Val Pro Ala Arg Thr Val Phe Gly Val Thr Thr
                155                 160                 165
Val Leu Thr Met Thr Thr Leu Ser Ile Ser Ala Arg Asn Ser Leu
                170                 175                 180
Pro Lys Val Ala Tyr Ala Thr Ala Met Asp Trp Phe Ile Ala Val
                185                 190                 195
Cys Tyr Ala Phe Val Phe Ser Ala Leu Ile Glu Phe Ala Thr Val
                200                 205                 210
Asn Tyr Phe Thr Lys Arg Gly Trp Ala Trp Asp Gly Lys Ser Val
                215                 220                 225
Val Asn Asp Lys Lys Lys Glu Lys Ala Ser Val Met Ile Gln Asn
                230                 235                 240
Asn Ala Tyr Ala Val Ala Val Ala Asn Tyr Ala Pro Asn Leu Ser
                245                 250                 255
Lys Asp Pro Val Leu Ser Thr Ile Ser Lys Ser Ala Thr Thr Pro
                260                 265                 270
Glu Pro Asn Lys Lys Pro Glu Asn Lys Pro Ala Glu Ala Lys Lys
                275                 280                 285
Thr Phe Asn Ser Val Ser Lys Ile Asp Arg Met Ser Arg Ile Val
                290                 295                 300
Phe Pro Val Leu Phe Gly Thr Phe Asn Leu Val Tyr Trp Ala Thr
                305                 310                 315
Tyr Leu Asn Arg Glu Pro Val Leu Gly Val Ser Pro
                320                 325
```

<210> 56
<211> 40
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510176CD1

<400> 56

```
Met Lys Phe Phe Ser Tyr Ile Leu Val Tyr Arg Arg Phe Leu Phe
1               5                   10                  15
```

Val Val Phe Thr Val Leu Val Leu Leu Pro Leu Pro Ile Val Leu
20 25 30
His Thr Lys Leu Ile Leu Thr Phe Pro Arg
35 40

<210> 57
<211> 104
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510541CD1

<400> 57
Met Glu Ala Pro Leu Gln Thr Glu Met Val Glu Leu Val Pro Asn
1 5 10 15
Gly Lys His Ser Glu Gly Leu Leu Pro Val Ile Thr Pro Met Ala
20 25 30
Gly Asn Gln Arg Val Glu Asp Pro Ala Arg Ser Cys Met Glu Gly
35 40 45
Lys Ser Phe Leu Gln Lys Ser Pro Ser Lys Glu Pro His Phe Thr
50 55 60
Asp Phe Glu Gly Lys Thr Ser Phe Gly Met Ser Val Phe Asn Leu
65 70 75
Ser Asn Ala Ile Met Gly Ser Gly Ile Leu Gly Leu Ala Tyr Ala
80 85 90
Met Ala Asn Thr Gly Ile Ile Leu Phe Leu His Pro Cys Leu
95 100

<210> 58
<211> 296
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510923CD1

<400> 58
Met Glu Ala Pro Leu Gln Thr Glu Met Val Glu Leu Val Pro Asn
1 5 10 15
Gly Lys His Ser Glu Gly Leu Leu Pro Val Ile Thr Pro Met Ala
20 25 30
Gly Asn Gln Arg Val Glu Asp Pro Ala Arg Ser Cys Met Glu Gly
35 40 45
Lys Ser Phe Leu Gln Lys Ser Pro Ser Lys Glu Pro His Phe Thr
50 55 60
Asp Phe Glu Gly Lys Thr Ser Phe Gly Met Ser Val Phe Asn Leu
65 70 75
Ser Asn Ala Ile Met Gly Ser Gly Ile Leu Gly Leu Ala Tyr Ala
80 85 90
Met Ala Asn Thr Gly Ile Ile Leu Phe Leu Phe Leu Leu Thr Ala
95 100 105
Val Ala Leu Leu Ser Ser Tyr Ser Ile His Leu Leu Leu Lys Ser
110 115 120

Ser Gly Val Val Gly Ile Arg Ala Tyr Glu Gln Leu Gly Tyr Arg
125 130 135
Ala Phe Gly Thr Pro Gly Lys Leu Ala Ala Ala Leu Ala Ile Thr
140 145 150
Leu Gln Asn Ile Gly Ala Met Ser Ser Tyr Leu Tyr Ile Ile Lys
155 160 165
Ser Glu Leu Pro Leu Val Ile Gln Thr Phe Leu Asn Leu Glu Glu
170 175 180
Lys Thr Ser Asp Trp Tyr Met Asn Gly Asn Tyr Leu Val Ile Leu
185 190 195
Val Ser Val Thr Ile Ile Leu Pro Leu Ala Leu Met Arg Gln Leu
200 205 210
Gly Tyr Leu Gly Tyr Ser Ser Gly Phe Ser Leu Ser Cys Met Val
215 220 225
Phe Phe Leu Ile Ala Val Ile Tyr Lys Lys Phe His Val Pro Cys
230 235 240
Pro Leu Pro Pro Asn Phe Asn Asn Thr Thr Gly Asn Phe Ser His
245 250 255
Val Glu Ile Val Lys Glu Lys Val Gln Leu Gln Val Glu Pro Glu
260 265 270
Ala Ser Ala Phe Cys Thr Pro Ser Tyr Phe Thr Leu Asn Ser Gln
275 280 285
Val Leu Thr Gly Gln Gly Lys Ala Gly Ala Gln
290 295

<210> 59
<211> 1364
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510984CD1

<400> 59
Met Pro Leu Ala Phe Cys Gly Ser Glu Asn His Ser Ala Ala Tyr
1 5 10 15
Arg Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Ala
20 25 30
Leu Asn Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro
35 40 45
Ile Leu Phe Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile
50 55 60
His His Ser Thr Trp Leu His Phe Pro Gly His Asn Leu Arg Trp
65 70 75
Ile Leu Thr Phe Met Leu Leu Phe Val Leu Val Cys Glu Ile Ala
80 85 90
Glu Gly Ile Leu Ser Asp Gly Val Thr Glu Ser His His Leu His
95 100 105
Leu Tyr Met Pro Ala Gly Met Ala Phe Met Ala Ala Val Ala Ser
110 115 120
Val Val Tyr Tyr His Asn Ile Glu Thr Ser Asn Phe Pro Lys Leu
125 130 135
Leu Ile Ala Leu Leu Val Tyr Trp Thr Leu Ala Phe Ile Thr Lys
140 145 150
Thr Ile Lys Phe Val Lys Phe Leu Asp His Ala Ile Gly Phe Ser

155 160 165
Gln Leu Arg Phe Cys Leu Thr Gly Leu Leu Val Ile Leu Tyr Gly
170 175 180
Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val Arg Arg Tyr
185 190 195
Ile Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu Asp Leu
200 205 210
Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu Leu
215 220 225
Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala
230 235 240
His Lys Lys Pro Ile Asp Leu Arg Ala Ile Gly Lys Leu Pro Ile
245 250 255
Ala Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Glu Ala Phe
260 265 270
Asp Ala Gln Val Arg Lys Asp Ile Gln Gly Thr Gln Gly Ala Arg
275 280 285
Ala Ile Trp Gln Ala Leu Ser His Ala Phe Gly Arg Arg Leu Val
290 295 300
Leu Ser Ser Thr Phe Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala
305 310 315
Gly Pro Leu Cys Ile Phe Gly Ile Val Asp His Leu Gly Lys Glu
320 325 330
Asn Asp Val Phe Gln Pro Lys Thr Gln Phe Leu Gly Val Tyr Phe
335 340 345
Val Ser Ser Gln Glu Phe Leu Ala Asn Ala Tyr Val Leu Ala Val
350 355 360
Leu Leu Phe Leu Ala Leu Leu Leu Gln Arg Thr Phe Leu Gln Ala
365 370 375
Ser Tyr Tyr Val Ala Ile Glu Thr Gly Ile Asn Leu Arg Gly Ala
380 385 390
Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met His Leu Ser Thr Ser
395 400 405
Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln Ile Cys Asn Leu
410 415 420
Val Ala Ile Asp Thr Asn Gln Leu Met Trp Phe Phe Phe Leu Cys
425 430 435
Pro Asn Leu Trp Ala Met Pro Val Gln Ile Ile Val Gly Val Ile
440 445 450
Leu Leu Tyr Tyr Ile Leu Gly Val Ser Ala Leu Ile Gly Ala Ala
455 460 465
Val Ile Ile Leu Leu Ala Pro Val Gln Tyr Phe Val Ala Thr Lys
470 475 480
Leu Ser Gln Ala Gln Arg Ser Thr Leu Glu Tyr Ser Asn Glu Arg
485 490 495
Leu Lys Gln Thr Asn Glu Met Leu Arg Gly Ile Lys Leu Leu Lys
500 505 510
Leu Tyr Ala Trp Glu Asn Ile Phe Arg Thr Arg Val Glu Thr Thr
515 520 525
Arg Arg Lys Glu Met Thr Ser Leu Arg Ala Phe Ala Ile Tyr Thr
530 535 540
Ser Ile Ser Ile Phe Met Asn Thr Ala Ile Pro Ile Ala Ala Val
545 550 555
Leu Ile Thr Phe Val Gly His Val Ser Phe Lys Glu Ala Asp
560 565 570
Phe Ser Pro Ser Val Ala Phe Ala Ser Leu Ser Leu Phe His Ile

575 580 585
Leu Val Thr Pro Leu Phe Leu Leu Ser Ser Val Val Arg Ser Thr
590 595 600
Val Lys Ala Leu Val Ser Val Gln Lys Leu Ser Glu Phe Leu Ser
605 610 615
Ser Ala Glu Ile Arg Glu Glu Gln Cys Ala Pro His Glu Pro Thr
620 625 630
Pro Gln Gly Pro Ala Ser Lys Tyr Gln Ala Val Pro Leu Arg Val
635 640 645
Val Asn Arg Lys Arg Pro Ala Arg Glu Asp Cys Arg Gly Leu Thr
650 655 660
Gly Pro Leu Gln Ser Leu Val Pro Ser Ala Asp Gly Asp Ala Asp
665 670 675
Asn Cys Cys Val Gln Ile Met Gly Gly Tyr Phe Thr Trp Thr Pro
680 685 690
Asp Gly Ile Pro Thr Leu Ser Asn Ile Thr Ile Arg Ile Pro Arg
695 700 705
Gly Gln Leu Thr Met Ile Val Gly Gln Val Gly Cys Gly Lys Ser
710 715 720
Ser Leu Leu Leu Ala Ala Leu Gly Glu Met Gln Lys Val Ser Gly
725 730 735
Ala Val Phe Trp Ser Ser Ser Leu Pro Asp Ser Glu Ile Gly Glu
740 745 750
Asp Pro Ser Pro Glu Arg Glu Thr Ala Thr Asp Leu Asp Ile Arg
755 760 765
Lys Arg Gly Pro Val Ala Tyr Ala Ser Gln Lys Pro Trp Leu Leu
770 775 780
Asn Ala Thr Val Glu Glu Asn Ile Ile Phe Glu Ser Pro Phe Asn
785 790 795
Lys Gln Arg Tyr Lys Met Val Ile Glu Ala Cys Ser Leu Gln Pro
800 805 810
Asp Ile Asp Ile Leu Pro His Gly Asp Gln Thr Gln Ile Gly Glu
815 820 825
Arg Gly Ile Asn Leu Ser Gly Gly Gln Arg Gln Arg Ile Ser Val
830 835 840
Ala Arg Ala Leu Tyr Gln His Ala Asn Val Val Phe Leu Asp Asp
845 850 855
Pro Phe Ser Ala Leu Asp Ile His Leu Ser Asp His Leu Met Gln
860 865 870
Ala Gly Ile Leu Glu Leu Leu Arg Asp Asp Lys Arg Thr Val Val
875 880 885
Leu Val Thr His Lys Leu Gln Tyr Leu Pro His Ala Asp Trp Ile
890 895 900
Ile Ala Met Lys Asp Gly Thr Ile Gln Arg Glu Gly Thr Leu Lys
905 910 915
Asp Phe Gln Arg Ser Glu Cys Gln Leu Phe Glu His Trp Lys Thr
920 925 930
Leu Met Asn Arg Gln Asp Gln Glu Leu Glu Lys Glu Thr Val Thr
935 940 945
Glu Arg Lys Ala Thr Glu Pro Pro Gln Gly Leu Ser Arg Ala Met
950 955 960
Ser Ser Arg Asp Gly Leu Leu Gln Asp Glu Glu Glu Glu Glu Glu
965 970 975
Glu Ala Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser Ser Met Leu
980 985 990
His Gln Arg Ala Glu Ile Pro Trp Arg Ala Cys Ala Lys Tyr Leu

```
                995                 1000                1005
Ser Ser Ala Gly Ile Leu Leu Leu Ser Leu Leu Val Phe Ser Gln
                1010                1015                1020
Leu Leu Lys His Met Val Leu Val Ala Ile Asp Tyr Trp Leu Ala
                1025                1030                1035
Lys Trp Thr Asp Ser Ala Leu Thr Leu Thr Pro Ala Ala Arg Asn
                1040                1045                1050
Cys Ser Leu Ser Gln Glu Cys Thr Leu Asp Gln Thr Val Tyr Ala
                1055                1060                1065
Met Val Phe Thr Val Leu Cys Ser Leu Gly Ile Val Leu Cys Leu
                1070                1075                1080
Val Thr Ser Val Thr Val Glu Trp Thr Gly Leu Lys Val Ala Lys
                1085                1090                1095
Arg Leu His Arg Ser Leu Leu Asn Arg Ile Ile Leu Ala Pro Met
                1100                1105                1110
Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser Ile Leu Asn Arg Phe
                1115                1120                1125
Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile Pro Ser Thr Leu
                1130                1135                1140
Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser Ala Leu Ala
                1145                1150                1155
Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu Leu Pro
                1160                1165                1170
Leu Ala Ile Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val Ala
                1175                1180                1185
Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu
                1190                1195                1200
Leu Ser His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg
                1205                1210                1215
Ala Phe Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr
                1220                1225                1230
Thr Asp Ser Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn
                1235                1240                1245
Arg Trp Leu Glu Val Arg Met Glu Tyr Ile Gly Ala Cys Val Val
                1250                1255                1260
Leu Ile Ala Ala Val Thr Ser Ile Ser Asn Ser Leu His Arg Glu
                1265                1270                1275
Leu Ser Ala Gly Leu Val Gly Leu Gly Leu Thr Tyr Ala Leu Met
                1280                1285                1290
Val Ser Asn Tyr Leu Asn Trp Met Val Arg Asn Leu Ala Asp Met
                1295                1300                1305
Glu Leu Gln Leu Gly Ala Val Lys Arg Ile His Gly Leu Leu Lys
                1310                1315                1320
Thr Glu Ala Glu Ser Tyr Glu Gly Leu Leu Gly Glu Arg Leu Arg
                1325                1330                1335
Glu Arg Gly Gly Glu Glu Ser Lys Glu Glu Cys Val Trp Val Gly
                1340                1345                1350
Gly His Lys Gly Ala Trp Gly Trp Gly Gly Thr Phe Gly Tyr
                1355                1360
```

<210> 60
<211> 895
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 7509332CB1

<400> 60
gagcacggag attcggcacg aggaggaaaa gcagcaaacc agatggagaa ggaaggccaa 60
aaagataagt agaaagctgg accatgagct ccttgagggc agggactgaa tggttactat 120
gtccccaggg cccagcatga ccttctcctg gattcctcat cttccttctg tgacctgtgt 180
ctccatcagt ttctcctccg gcatcttttt cttacaggat tcttacctca ggaatagatg 240
gacatggcct ggcagatgat gcagctgctg cttctggctt tggtgactgc tgcgggggagt 300
gcccagccca ggagtgcgcg ggccaggacg gacctgctca atgtctgcat gaacgccaag 360
caccacaaga cacagcccag ccccgaggac gagctgtatg gccagtgcag tccctggaag 420
aagaatgcct gctgcacggc cagcaccagc caggagctgc acaaggacac ctcccgcctg 480
tacaacttta actgggatca ctgtgagcgc tggtgggagg actgtcgcac ctcctacacc 540
tgcaaaagca actggcacaa aggctggaat tggacctcag ggattaatga gtgtccggcc 600
ggggccctct gcagcacctt tgagtcctac ttccccactc cagccgccct ttgtgaaggc 660
ctctggagcc actccttcaa ggtcagcaac tatagtcgag ggagcggccg ctgcatccag 720
atgtggtttg actcagccca gggcaacccc aatgaggagg tggccaagtt ctatgctgcg 780
gccatgaatg ctggggcccc gtctcgtggg attattgatt cctgatccaa gaagggtcct 840
ctggggttct tccaacaacc tattctaata gacaaatcca catgaaaaaa aaaaa 895

<210> 61
<211> 1623
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509102CB1

<400> 61
cagcaaagcc ctgacctgac ctcctgatgc tcaggagaag ccatgggccc ctcctgtcct 60
gtgttcctgt ccttcacaaa gctcggcctg tggtggctcc ttctgacccc agcaggtgga 120
gaggaagcta agcgcccacc tcccagggct cctggagacc cactctcctc tcccagtccc 180
acggcattgc cgcagggagg ctcgcatacc gagactgagg accggctctt caaacacctc 240
ttccggggct acaaccgctg ggcgcgcccg gtgcccaaca cttcagacgt ggatgagaag 300
aaccaaatga tgaccaccaa cgtctggcta aaacaggagt ggagcgacta caaactgcgc 360
tggaacccca ctgattttgg caacatcaca tctctcaggg tcccttctga gatgatctgg 420
atccccgaca ttgttctcta caacaaaact gcaagatgaa gtttggctcc tggacttatg 480
acaaggccaa gatcgacctg gagcagatgg agcagactgt ggacctgaag gactactggg 540
agagcggcga gtgggccatc gtcaatgcca cgggcaccta caacagcaag aagtacgact 600
gctgcgccga gatctacccc gacgtcacct acgccttcgt catccggcgg ctgccgctct 660
tctacaccat caacctcatc atcccctgcc tgctcatctc ctgcctcact gtgctggtct 720
tctacctgcc ctccgactgc ggcgagaaga tcacgctgtg catttcggtg ctgctgtcac 780
tcaccgtctt cctgctgctc atcactgaga tcatcccgtc cacctcgctg gtcatcccgc 840
tcatcggcga gtacctgctg ttcaccatga tcttcgtcac cctgtccatc gtcatcaccg 900
tcttcgtgct caatgtgcac caccgctccc ccagcaccca caccatgccc cactgggtgc 960
ggggggccct tctgggctgt gtgccccggt ggcttctgat gaaccggccc ccaccacccg 1020
tggagctctg ccacccccta cgcctgaagc tcagcccctc ttatcactgg ctggagagca 1080
acgtggatgc cgaggagagg gaggtggtgg tggaggagga ggacagatgg gcatgtgcag 1140
gtcatgtggc cccctctgtg ggcaccctct gcagccacgg ccacctgcac tctgggggcct 1200
caggtcccaa ggctgaggct ctgctgcagg agggtgagct gctgctatca ccccacatgc 1260
agaaggcact ggaaggtgtg cactacattg ccgaccacct gcggtctgag gatgctgact 1320
cttcggtgaa ggaggactgg aagtatgttg ccatggtcat cgacaggatc ttcctctggc 1380
tgtttatcat cgtctgcttc ctggggacca tcggcctctt tctgcctccg ttcctagctg 1440
gaatgatctg actgcacctc cctcgagctg gctcccaggg caaaggggag ggttcttgga 1500
tgtggaaggg ctttgaacaa tgtttagatt tggagatgag cccaaagtgc cagggagaac 1560

agccaggtga ggtgggaggt tggagagcca ggtgaggtct ctctaagtca ggctggggtt 1620
gaa 1623

<210> 62
<211> 1802
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509132CB1

<400> 62
gaatcagctg tgacgagccg catcagctta tactggcgca gtgtgctgga atgcgtcctt 60
caggatcggt tctttcatct tcgccgcccc tgcgcgtcca gctcttctaa gacgagatgc 120
cgtcgggctt ccaacagata ggctccgaag atggggaacc ccctcagcag cgagtgactg 180
ggaccctggt ccttgctgtg ttctctgcgg tgcttggctc cctgcagttt gggtacaaca 240
ttggggtcat caatgcccct cagaaggtga ttgaacagag ctacaatgag acgtggctgg 300
ggaggcaggg gcctgaggga cccagctcca tccctccagg caccctcacc accctctggg 360
ccctctccgt ggccatcttt tccgtgggcg gcatgatttc ctccttcctc attggtatca 420
tctctcagtg gcttggaagg aaaagggcca tgctggtcaa caatgtcctg gcggtgctgg 480
ggggcagcct catgggcctg gccaatgctg ctgcctccta tgaaatgctc atccttggac 540
gattcctcat tggcgcctac tcaggtgctg ggcttggagt ccctcctggg cactgccagc 600
ctgtggccac tgctcctggg cctcacagtg ctacctgccc tcctgcagct ggtcctgctg 660
cccttctgtc ccgagagccc ccgctacctc tacatcatcc agaatctcga ggggcctgcc 720
agaaagagtc tgaagcgcct gacaggctgg gccgatgttt ctggagtgct ggctgagctg 780
aaggatgaga agcggaagct ggagcgtgag cggccactgt ccctgctcca gctcctgggc 840
agccgtaccc accggcagcc cctgatcatt gcggtcgtgc tgcagctgag ccagcagctc 900
tctggcatca atgctgtttt ctattattcg accagcatct tcgagacagc aggggtaggc 960
cagcctgcct atgccaccat aggagctggt gtggtcaaca cagtcttcac cttggtctcg 1020
gtgttgttgg tggagcgggc ggggcgccgg acgctccatc tcctgggcct ggcgggcatg 1080
tgtggctgtg ccatcctgat gactgtggct ctgctcctgc tggagcgagt tccagccatg 1140
agctacgtct ccattgtggc catctttggc ttcgtggcat tttttgagat tggccctggc 1200
cccattcctt ggttcatcgt ggccgagctc ttcagccagg gaccccgccc ggcagccatg 1260
gctgtggctg gtttctccaa ctggacgagc aacttcatca ttggcatggg tttccagtat 1320
gttgcggagg ctatggggcc ctacgtcttc cttctatttg cggtcctcct gctgggcttc 1380
ttcatcttca ccttcttaag agtacctgaa actcgaggcc ggacgtttga ccagatctca 1440
gctgccttcc accggacacc ctctctttta gagcaggagg tgaaacccag cacagaactt 1500
gagtatttag ggccagatga gaacgactga ggggccaggc aggggtggga gagccagctc 1560
tctctacccg gcccagagac cccttccttt cctctgcagc actttaaccc tctcttccct 1620
attatttccg ggtggaaaag aatccctgca gcctggtaga attgggaagc tgggggaagg 1680
gtggtctgag cacccccctca ttcccctcgt gtgactctct tggattattt atgtgttgtg 1740
gtttggccgt ggccatcagg gtgggccact ctcccctccc tcttccttcc cccatcccct 1800
tt 1802

<210> 63
<211> 2139
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509136CB1

<400> 63
agcctttggt gaagattgga tcgaaatcag accaatggac aagctctggc cgtgggtggt 60

ggacgggcct ggagtagaac tggaacgaga aggatgaaga gatgagcaca aaggtgtact 120
tagacctgga gtggactgac tacaggctga gctgggaccc tgcggagcac gacggcatcg 180
attcgctccg catcacggcg gaatccgtgt ggctccctga cgtggtgcta ctgaacaaca 240
atgatgggaa ttttgacgtg gctctggaca ttagcgtcgt ggtgtcctcc gacggctccg 300
tgcgttggca acccccgggc atctatcgca gcagctgcag catccaggtc acctacttcc 360
ccttcgactg gcagaattgc actatggtgt tcagctccta cagctacgac agctcggagg 420
tcagcctgca gacaggcctg ggtcctgacg ggcaagggca tcaggaaatc cacattcatg 480
aagggacttt cattgagaat ggccagtggg agattatcca caagccctct cggctaatcc 540
agcctccagg cgatcctagg ggagggaggg aaggacagcg ccaggaagtc atcttctacc 600
tcatcatccg ccgcaagcct ctcttctacc tggtcaacgt cattgcccca tgcatcctca 660
tcactcttct ggccatcttc gtcttctacc tgccaccaga tgcagtcatc cttagtgtcg 720
tggttctcaa cctgcaccac cgctcacccc acacccacca aatgcccctt tgggtccgtc 780
agatcttcat tcacaaactt ccgctgtacc tgcgtctaaa aaggcccaaa cccgagagag 840
acctgatgcc ggagcccccт cactgttctt ctccaggaag tggctggggt cggggaacag 900
atgaatattt catccggaag ccgccaagtg attttctctt ccccaaaccc aataggttcc 960
agcctgaact gtctgcccct gatctgcggc gatttatcga tggtccaaac cgggctgtgg 1020
ccctgcttcc ggagctacgg gaggtcgtct cctctatcag ctacatcgct cgacagctgc 1080
aggaacagga ggaccacgat gcgctgaagg aggactggca gtttgtggcc atggtagtgg 1140
accgcctctt cctgtggact ttcatcatct tcaccagcgt tgggaccctа gtcatcttcc 1200
tggacgccac gtaccacttg ccccctccag accccttttcc ttgaagactg gagggttgag 1260
acccaggccc cctgccagtt gaagtgagag tttggtgata ctgtcaagcc ccatccttct 1320
ctgcctctta actccttcac gaggaatctg ggcctcttat ttcgtttctg ggactgcat 1380
tggactgagg gctgggtagg caggtgtctt ggacccacct gaaatgcagt atcatctgat 1440
ttactctttg ggatcttgaa gaagctcttt tgggtatcaa cacctaggtc gccagtgaaa 1500
tagaacacag aacaggaact agattataag ccttatgagg tcaagaaatg tgacttggcc 1560
gggcgcggtg gctcacgcct gtaatcccag cactttggga ggccaaggcg ggcggatcac 1620
ctgaggtcgg gagtttgaga ccagcccgac caacatggag aaaccctgtc tctactaaaa 1680
atacaaaatt agccaggtgt ggtggtacat gcctgtaatc ccagctacta gggaggctga 1740
ggcaggagaa tcactcgaac ccgggaggca gaggttgcag tgagtcaaga tcacgccatt 1800
gcactccagc ctgggcaaca agagcgaaac tccatctcaa aaaaaaaaaa gtgtcttgtt 1860
cactgcagca tgcccagtgc cacgcacagg tgctgaccta tagtaagtgc ttaacatttg 1920
ttgaataggg gaaagaaatt tccggaagta aatacagcaa ttaataatgt tttataagctg 1980
ggcatggtcg ctcaggccag taatcccagt gacttaggag gctgaggtgg gaggattact 2040
tgaaacccac agtttgagac cagcctgggc aacatagtga gaccctgtct ctaaaaaaat 2100
aaaaatagaa ataaagtagt gcttattgtt tgcaaaaaa 2139

<210> 64
<211> 1461
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509178CB1

<400> 64
cacaagctcc ggtagcccat ggagccctgg cctctcctcc tgctctttag cctttgctca 60
gctggcctcg tcctgggctc cgaacatgag acccgtctgg tggcaaagct atttaaagac 120
tacagcagcg tggtgcggcc agtggaagac caccgccagg tcgtggaggt caccgtgggc 180
ctgcagctga tacagctcat caatgtggat gaagtaaatc agatcgtgac aaccaatgtg 240
cgtctgaaac agaactgcag catgaagctg ggcacctgga cctacgacgg ctctgtcgtg 300
gccatcaacc cggaaagcga ccagccagac ctgagcaact tcatggagag cggggagtgg 360
gtgatcaagg agtcccgggg ctggaagcac tccgtgacct attcctgctg ccccgacacc 420
ccctacctgg acatcaccta ccacttcgtc atgcagcgcc tgcccctcta cttcatcgtc 480
aacgtcatca tcccctgcct gctcttctcc ttcttaactg gcctggtatt ctacctgccc 540
acagactcag gggagaagat gactctgagc atctctgtct tactgtcttt gactgtgttc 600

cttctggtca tcgtggagct gatcccctcc acgtccagtg ctgtgccctt gattggaaaa 660
tacatgctgt tcaccatggt gttcgtcatt gcctccatca tcatcactgt catcgtcatc 720
aacacacacc accgctcacc cagcacccat gtcatgccca actgggtgcg gaaggttttt 780
atcgacacta tcccaaatat catgtttttc tccacaatga aaagaccatc cagagaaaag 840
caagacaaaa agatttttac agaagacatt gatatctctg acatttctgg aaagccaggg 900
cctccaccca tgggcttcca ctctcccctg atcaaacacc ccgaggtgaa aagtgccatc 960
gagggcatca agtacatcgc agagaccatg aagtcagacc aggagtctaa caatgcggcg 1020
gcagagtgga agtacgttgc aatggtgatg gaccacatac tcctcggagt cttcatgctt 1080
gtttgcatca tcggaaccct agccgtgttt gcaggtcgac tcattgaatt aaatcagcaa 1140
ggatgagcag aaaatgagct gagcttagct ctgccctgga acctaccaga gcagagaagg 1200
gcaggagagg aagatttgtc tacttgctcc actcgcactt atcaaacgtg ttatattcca 1260
tacttattat tgatgataag atttaccttt atgtaagttt atggccttga agtgttttca 1320
tattgcttct ccctttagtt ctgctgtctc cctgaagagt gaaccctctt tagtaaatga 1380
aactaatcac taaaaaaagt gttcatttcc agtgtctgga agagtttttg ccaggataac 1440
cgaggtttct gttgcattgc a 1461

<210> 65
<211> 738
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509214CB1

<400> 65
gacagcggcc tggctaactc ctgccaggca gtgcccttcc cggagcgtgc cctcgccgct 60
gagctcccct gaacagcagc tgcagcagcc atggccccgc cctgggtgcc cgccatgggc 120
ttcacgctgg cgcccagcca tggggtacgg ctcctacctg gtctggaaag agctgggagg 180
cttcacagag aaggctgtgg ttcccctggg cctctacact gggcagctgg ccctgaactg 240
ggcatggccc cccatcttct ttggtgcccg acaaatgggc tgggccttgg tggatctcct 300
gctggtcagt ggggcggcgg cagccactac cgtggcctgg taccaggtga gcccgctggc 360
cgcccgcctg ctctacccct acctggcctg gctggccttc gcgaccacac tcaactactg 420
cgtatggcgg gacaaccatg gctggcgtgg gggacggcgg ctgccagagt gagtgcccgg 480
cccaccaggg actgcagctg caccagcagg tgccatcacg cttgtgatgt ggtggccgtc 540
acgctttcat gaccactggg cctgctagtc tgtcagggcc ttggcccagg ggtcagcaga 600
gcttcagagg tggccccacc tgagccccca cccgggagca gtgtcctgtg ctttctgcat 660
gcttagagca tgttcttgga acatggaatt ttataagctg aataaagttt ttgacttcct 720
ttaaaaaaaa aaaaaaaa 738

<210> 66
<211> 2106
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509244CB1

<400> 66
agcctttggt gaagattgga tcgaaatcag accaatggac aagctctggc cgtgggtggt 60
ggacgggcct ggagtagaac tggaacgaga aggatgaaga gatgagcaca aaggtgtact 120
tagacctgga gtggactgac tacaggctga gctgggaccc tgcggagcac gacggcatcg 180
attcgctccg catcacggcg gaatccgtgt ggctccctga cgtggtgcta ctgaacaaca 240
atgatgggaa ttttgacgtg gctctggaca ttagcgtcgt ggtgtcctcc gacggctccg 300
tgcgttggca acccccgggc atctatcgca gcagctgcag catccaggtc acctacttcc 360

ccttcgactg gcagaattgc actatggtgt tcagctccta cagctacgac agctcggagg 420
tcagcctgca gacaggcctg ggtcctgacg ggcaagggca tcaggaaatc cacattcatg 480
aagggacttt cattgagaat ggccagtggg agattatcca caagccctct cggctaatcc 540
agcctccagg cgatcctagg ggagggaggg aaggacagcg ccaggaagtc atcttctacc 600
tcatcatccg ccgcaagcct ctcttctacc tggtcaacgt cattgcccca tgcatcctca 660
tcactcttct ggccatcttc gtcttctacc tgccaccaga tgcaggagag aagatggggc 720
tctcaatctt tgccctgctg acccttactg tgttcctgct gctgctggct gacaaagtac 780
ctgagacctc actatcagta cccattatta tcaagtacct catgtttacc atggtcctcg 840
tcaccttctc agtcatcctt agtgtcgtgg ttctcaacct gcaccaccgc tcaccccaca 900
cccaccaaat gcccctttgg gtccgtcaga tcttcattca caaacttccg ctgtacctgc 960
gtctaaaaag gcccaaaccc gagagagacc tgatgccgga gctacgggag gtcgtctcct 1020
ctatcagcta catcgctcga cagctgcagg aacaggagga ccacgatgcg ctgaaggagg 1080
actggcagtt tgtggccatg gtagtggacc gcctcttcct gtggactttc atcatcttca 1140
ccagcgttgg gaccctagtc atcttcctgg acgccacgta ccacttgccc cctccagacc 1200
cctttccttg aagactggag ggttgagacc caggcccccct gccagttgaa gtgagagttt 1260
ggtgatactg tcaagcccca tccttctctg cctcttaact ccttcacgag gaatctgggc 1320
ctcttatttc gtttctgggg actgcattgg actgagggct gggtaggcag gtgtcttgga 1380
cccacctgaa atgcagtatc atctgattta ctctttggga tcttgaagaa gctcttttgg 1440
gtatcaacac ctaggtcgcc agtgaaatag aacacagaac aggaactaga ttataagcct 1500
tatgaggtca agaaatgtga cttggccggg cgcggtggct cacgcctgta atcccagcac 1560
tttgggaggc caaggcgggc ggatcacctg aggtcgggag tttgagacca gcccgaccaa 1620
catggagaaa ccctgtctct actaaaaata caaaattagc caggtgtggt ggtacatgcc 1680
tgtaatccca gctactaggg aggctgaggc aggagaatca ctcgaacccg ggaggcagag 1740
gttgcagtga gtcaagatca cgccattgca ctccagcctg ggcaacaaga gcgaaactcc 1800
atctcaaaaa aaaaaaagtg tcttgttcac tgcagcatgc ccagtgccac gcacaggtgc 1860
tgacctatag taagtgctta acatttgttg aataggggaa agaaatttcc ggaagtaaat 1920
acagcaatta ataatgttta taagctgggc atggtcgctc aggccagtaa tcccagtgac 1980
ttaggaggct gaggtgggag gattacttga aacccacagt ttgagaccag cctgggcaac 2040
atagtgagac cctgtctcta aaaaaataaa aatagaaata aagtagtgct tattgtttgc 2100
aaaaaa 2106

<210> 67
<211> 2334
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509256CB1

<400> 67
ctcgagccgc tcgagccgcg ggcggctggg acgcagctgc cccgctcgg cgcccgctgc 60
acccttagca gccactgcca cctgggcccg gagcctccac gatctcgccc ggcgattgtg 120
ggcagggggcg cctccggatc gatcttctga aattcaagtt ttcaagatga agtttttatt 180
gacaactgcc tttttaattt taatttcctt gtgggtggaa gaagcctatt ctaaggaaaa 240
gtcttcaaag aaagggaagg ggaaaaagaa gcagtatcta tgcccatctc agcagtcagc 300
agaggacctt gcccgagtac ctgccaactc cactagcaat atcttgaaca ggttattggt 360
cagttatgat cccaggataa gaccaaactt caaaggcatt cctgttgatg tagtagtcaa 420
catttttatt aacagttttg gatccattca agaaacaaca atggactata gagttaacat 480
cttcctgaga caaaaatgga atgaccccag gctgaagctc cccagtgatt ttaggggttc 540
agatgcactg acagtggatc caacaatgta caagtgttta tggaaacctg atttattttt 600
tgcaaatgaa aaaagtgcca attttcatga tgtgacccag gaaaacatcc tcctctttat 660
ttttcgtgat ggagatgtcc ttgtcagcat gaggttatct attactcttt catgcccttt 720
ggacttgaca ttgtttccca tggatacaca acgttgcaag atgcaactgg agagctttgg 780
ttacacaact gatgatttac gatttatctg gcagtcagga gatcctgtgc aattagaaaa 840
aattgccttg cctcaatttg atatcaaaaa ggaagatatt gaatatggta actgtacaaa 900

atactataaa ggcacgggct actacacatg cgtggaagtc atcttcaccc tgaggaggca 960
ggtcggcttt tacatgatgg gggtctacgc cccaaccctg ctcattgttg ttctctcctg 1020
gctttccttc tggatcaacc cggacgcgag tgctgccaga gtgcccctgg gttggtgaga 1080
ccagatgcaa aaaagtttgt acttctaagt ctgatctgag atctaatgac ttcagcattg 1140
ttggaagctt accaagagat tttgaactat ccaattatga ctgctatgga aaacccattg 1200
aagttaacaa cggacttggg aaatctcagg ctaagaacaa caagaagcct cccccctgcga 1260
aacctgttat tccaacagca gcaaagcgaa ttgatcttta tgcaagagca ttgtttcctt 1320
tctgcttctt gttcttcaat gttatatatt ggtctatata tttatgataa atcttttcca 1380
tttgtacaaa ataaaattcc atttcattgt gacctactcc tttcataaat gccaatctgt 1440
gagaactttt gaattttcat agcaacattg cattttggat gccatttgat tgtaataaaa 1500
ctgtggcacc ttaattttga atggcagcat gatcatgtaa tatctgtgct ctaataacga 1560
tgtatatatg tatagtgaac atattgctta gtaacaaatg aaggacaagc atactacata 1620
atataatcca tacaattctc ttcagttagt gtaaactgca aatactacag ataattctga 1680
taataaaatg atatgcacgc tgaatcctgc tatggtcacc attctaatgt atgtagtatt 1740
tcaaatttcc ttccttgtaa ctttcaaaga aagccatctt attcttgtaa aattttagat 1800
ggtattatca cagatttaaa aaggttgtat tacatattgt ttaaactttg taagtagaaa 1860
tatatctgtt ataattatac aggctctgtg gagaaataaa gttcaaaaaa tattaatttg 1920
taaaatcagc tcgttttaaa gtgtgcttgt gttgtcaaaa atatcagata gtaatacaca 1980
gtgagcattt ttaaacaaag ggaaacctat atttatgtaa ctgtatactg aattctgaca 2040
aaataaaaaa agatacctta ttgacgaaat atttaggata aacaaaattc tatttaatcc 2100
accttaaaac ctaaatgtat tttcatggat ttcatttgtt ggtacatatt acacaaaaca 2160
ttgtgcctta aaatgagtca tacatctttt aaattggaat gcagtaatag atatgtgatt 2220
ttacatcatt tttaagaaac caaggggaag taataagttg aaaaagaaat ccataactat 2280
taaaagattt taactttttt attttattaa aatgcttgca tattttaagt aaaa 2334

<210> 68
<211> 1475
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509395CB1

<400> 68
ccacggccag tgtgctcgaa tcggtacttc acaagctccg gtagcccatg gagccctggc 60
ctctcctcct gctctttagc ctttgctcag ctggcctcgt cctgggctcc gaacatgaga 120
cccgtctggt ggcaaagcta tttaaagact acagcagcgt ggtgcggcca gtggaagacc 180
accgccaggt cgtggaggtc accgtgggcc tgcagctgat acagctcatc aatgtggatg 240
aagtaaatca gatcgtgaca accaataact gcagcatgaa gctgggcacc tggacctacg 300
acggctctgt cgtggccatc aacccggaaa gcgaccagcc agacctgagc aacttcatgg 360
agagcgggga gtgggtgatc aaggagtccc ggggctggaa gcactccgtg acctattcct 420
gctgccccga cacccccctac ctggacatca cctaccactt cgtcatgcag cgcctgcccc 480
tctacttcat cgtcaacgtc atcatcccct gcctgctctt ctccttctta actggcctgg 540
tattctacct gcccacagac tcaggggaga agatgactct gagcatctct gtcttactgt 600
ctttgactgt gttccttctg gtcatcgtgg agctgatccc ctccacgtcc agtgctgtgc 660
ccttgattgg aaaatacatg ctgttcacca tggtgttcgt cattgcctcc atcatcatca 720
ctgtcatcgt catcaacaca caccaccgct cacccagcac ccatgtcatg cccaactggg 780
tgcggaaggt ttttatcgac actatcccaa atatcatgtt tttctccaca atgaaaagac 840
catccagaga aaagcaagac aaaaagattt ttacagaaga cattgatatc tctgacattt 900
ctggaaagcc agggcctcca cccatgggct tccactctcc cctgatcaaa caccccgagg 960
tgaaaagtgc catcgagggc atcaagtaca tcgcagagac catgaagtca gaccaggagt 1020
ctaacaatgc ggcggcagag tggaagtacg ttgcaatggt gatggaccac atactcctcg 1080
gagtcttcat gcttgtttgc atcatcggaa ccctagccgt gtttgcaggt cgactcattg 1140
aattaaatca gcaaggatga gcagaaaatg agctgagctt agctctgccc tggaacctac 1200
cagagcagag aagggcagga gaggaagatt tgtctacttg ctccactcgc acttatcaaa 1260

cgtgttatat tccatactta ttattgatga taagatttac ctttatgtaa gtttatggcc 1320
ttgaagtgtt ttcatattgc ttctcccttt agttctgctg tctccctgaa gagtgaaccc 1380
tctttagtaa atgaaactaa tcactaaaaa aagtgttcat ttccagtgtc tggaagagtt 1440
tttgccagga taaccgaggt ttctgttgca ttgca 1475

<210> 69
<211> 1295
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503287CB1

<400> 69
ccaggacccg ccgccggctg ccggcttgcc gaagccccct caggatcccc tcaacaagga 60
tggaactgaa ggccgaggag gaggaggtgg gtggcgtcca gccggtgagc atccaggcct 120
tcgccagcag ctccacactg cacggcctgg cccacatctt ctcctacgag cggctgtctc 180
tgaagcgggc actgtgggcc ctgtgcttcc tgggctcgct ggctgtgctg ctgtgtgtgt 240
gcacggagcg tgtgcagtac tacttccact accaccatgt caccaagctc gacgaggtgg 300
ctgcctctca gcttaccttc cctgctgtca cgctgtgcaa cctcaacgag ttccgcttta 360
gccaagtctc caagaatgac ctgtatcatg ctggggagct gctggccctg ctcaacaaca 420
ggtatgagat accagacaca cagatggcag atgaaaagca gctggagata ctgcaggaca 480
aagccaactt ccgcagcttc aaacccaaac ccttcaacat gcgtgagttc tacgaccgag 540
ctgggcacga cattcgagac atgctgctct cctgccactt ccggggggag gtctgcagcg 600
ctgaagactt caaggtggtc ttcacacgct atggaaagtg ctacacgttc aactcgggcc 660
gagatgggcg gccgcggctg aagaccatga agggtgggac gggcaatggg ctggaaatca 720
tgctggacat ccagcaggac gagtacctgc ctgtgtgggg ggagactgac gagacgtcct 780
tcgaagcagg catcaaagtg cagatctttc ctcttgtctg tggtaaggaa ggagtcttga 840
ccatagagtc ctctctctgc ctctatccca ttctttttac atttaacaaa actaatctaa 900
aaaagaacta aaaagggaga acggggcaag ggacctcagg ctgcccctct ctcctccatg 960
ctgcctcccc tagctcccag cctgaattct gtctatctag ctgtctgcca tctgagtgtc 1020
catctacatt ctgctgccac cagtcaccaa aggcccttcc cagtgagggg tggaagggat 1080
ctctggggtc tggaatttgg ccccaaacca gagaatgtac cttaaggggg agggctagtg 1140
tgggggaggg aggcttcccc agccttaaga gaccctctca gcccagtgac tgtccccaaa 1200
cccaagtctc ctggcaggaa ctaaaacctc agccccactc tctcacacca tgtggaatct 1260
cgtgggggtc ggggatcccc ttaagaagtg gtaat 1295

<210> 70
<211> 1386
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503320CB1

<400> 70
tgcagctccg ggactcaaca tgcgctgctc gccgggaggc gtctggctgg cgctggccgc 60
gtcgctcctg cacgtgtccc tgcaaggcga gttccagagg aagctttaca aggagctggt 120
caagaactac aatcccttgg agaggcccgt ggccaatgac tcgcaaccac tcaccgtcta 180
cttctccctg agcctcctgc agatcatgga cgtggatgag aagaaccaag ttttaaccac 240
cacgaccccg acgggggcaa gatgcccaag tggaccagag tcatccttct gaactggtgc 300
gcgtggttcc tgcgaatgaa gaggcccggg gaggacaagg tgcgcccggc ctgccagcac 360
aagcagcggc gctgcagcct ggccagtgtg gagatgagcg ccgtggcgcc gccgcccgcc 420
agcaacggga acctgctgta catcggcttc cgcggcctgg acggcgtgca ctgtgtcccg 480

acccccgact ctggggtagt gtgtggccgc atggcctgct cccccacgca cgatgagcac 540
ctcctgcacg gcgggcaacc ccccgagggg gacccggact tggccaagat cctggaggag 600
gtccgctaca ttgccaaccg cttccgctgc caggacgaaa gcgaggcggt ctgcagcgag 660
tggaagttcg ccgcctgtgt ggtggaccgc ctgtgcctca tggccttctc ggtcttcacc 720
atcatctgca ccatcggcat cctgatgtcg gctcccaact tcgtggaggc cgtgtccaaa 780
gactttgcgt aaccacgcct ggttctgtac atgtggaaaa ctcacagatg ggcaaggcct 840
ttggcttggc gagatttggg ggtgctaatc caggacagca ttacacgcca caactccagt 900
gttcccttct ggctgtcagt cgtgttgctt acggtttctt tgttacttta ggtagtagaa 960
tctcagcact ttgtttcata ttctcagatg ggctgataga tatccttggc acatccgtac 1020
catcggtcag cagggccact gagtagtcat tttgcccatt agcccactgc ctggaaagcc 1080
cttcggagag ctccccatgg ctcctcacca ccgagacagt tggttttgca tgtctgcatg 1140
aaggtctacc tgaaaattca acatttgctt tttgcttgtg tacaaaccca gattgaagct 1200
aaaataaacc agactcacta aatcctttcc aataattgac tggtggaagg aaaacaaaaa 1260
acaaaaacta aaaacctctt agcttttctg caattcaact ttttattttt atttttattt 1320
ctatcaaaga cggtagagag aaacagcttg atgctgtttc tacattaaaa aaaaaattgg 1380
tcggtc 1386

<210> 71
<211> 2213
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503335CB1

<400> 71
gttactggga gggggcttgc tgtggccctg tcaggaagag tagagctctg gtccagctcc 60
gcgcagggag ggaggctgtc accatgccgg cctgctgcag ctgcagtgat gttttccagt 120
atgagacgaa caaagtcact cggatccaga gcatgaatta tggcaccatt aagtggttct 180
tccacgtgat catcttttcc tacgtttgct ttgctctggt gagtgacaag ctgtaccagc 240
ggaaagagcc tgtcatcagt tctgtgcaca ccaaggtgaa ggggatagca gaggtgaaag 300
aggagatcgt ggagaatgga gtgaagaagt tggtgcacag tgtctttgac accgcagact 360
acaccttccc tttgcagggg aactctttct tcgtgatgac aaactttctc aaaacagaag 420
gccaagagca gcggttgtgt cccgagtatc ccacccgcag gacgctctgt tcctctgacc 480
gaggttgtaa aaagggatgg atggacccgc agagcaaagg aattcagacc ggaaggtgtg 540
tagtgcatga agggaaccag aagacctgtg aagtctctgc ctggtgcccc atcgaggcag 600
tggaagaggc cccccggcct gctctcttga acagtgccga aaacttcact gtgctcatca 660
agaacaatat cgacttcccc ggccacaact acaccacgag aaacatcctg ccaggtttaa 720
acatcacttg taccttccac aagactcaga atccacagtg tcccattttc cgactaggag 780
acatcttccg agaacaggcg ataatttttc agatgtggca attcagatac gccaagtact 840
acaaggaaaa caatgttgag aaacggactc tgataaaagt cttcgggatc cgttttgaca 900
tcctggtttt tggcaccgga ggaaaatttg acattatcca gctggttgtg tacatcggct 960
caaccctctc ctacttcggt ctggccgctg tgttcatcga cttcctcatc gacacttact 1020
ccagtaactg ctgtcgctcc catatttatc cctggtgcaa gtgctgtcag ccctgtgtgg 1080
tcaacgaata ctactacagg aagaagtgcg agtccattgt ggagccaaag ccgacattaa 1140
agtatgtgtc ctttgtggat gaatcccaca ttaggatggt gaaccagcag ctactaggga 1200
gaagtctgca agatgtcaag ggccaagaag tcccaagacc tgcgatggac ttcacagatt 1260
tgtccaggct gcccctggcc ctccatgaca cacccccgat tcctggacaa ccagaggaga 1320
tacagctgct tagaaaggag gcgactccta gatccaggga tagccccgtc tggtgccagt 1380
gtggaagctg cctcccatct caactccctg agagccacag gtgcctggag gagctgtgct 1440
gccggaaaaa gccgggggcc tgcatcacca cctcagagct gttcaggaag ctggtcctgt 1500
ccagacacgt cctgcagttc ctcctgctct accaggagcc cttgctggcg ctggatgtgg 1560
attccaccaa cagccggctg cggcactgtg cctacaggtg ctacgccacc tggcgcttcg 1620
gctcccagga catggctgac tttgccatcc tgcccagctg ctgccgctgg aggatccgga 1680
aagagtttcc aaagagtgaa gggcagtaca gtggcttcaa gagtccttac tgaagccagg 1740

caccgtggct cacgtctgta atcccagcgc tttgggaggc cgaggcaggc agatcacctg 1800
aggtcgggag ttggagaccc gcctggctaa caaggcgaaa tcctgtctgt actaaaaata 1860
caaaaatcag ccagacatgg tggcatgcac ctgcaatccc agctactcgg gaggctgagg 1920
cacaagaatc acttgaaccc gggaggcaga ggttgtagtg agcccagatt gtgccactgc 1980
tctccagcct gggaggcaca gcaaactgtc ccaaaaaaaa caaaagagtc cttaccaata 2040
gccgggggct tgcagtaccc atgttaactt gaccatttac caagcaattg aacttcacct 2100
gcaaagctct gtggccactt ttccagccaa agggaaatat gcttcacctc ggtggccctc 2160
ggggtctgaa gcaaaggggc ctggttaaac aaccaaaaat ccttaaatgg agc 2213

<210> 72
<211> 1289
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503952CB1

<400> 72
gagaaattga gcggcattcc atctggtagg caagtttgca tttctccttt ttgggatctg 60
cccaggaatg ttgtcaagtg taatggctcc cctgtgggcc tgcatcctgg tggctgcagg 120
aattctagcc acagatacac atcatcccca ggattctgct ctgtatcatc tcagcaagca 180
gctattacag aaatatcata aagaagtgag acctgtttac aactggacca aggccaccac 240
agtctacctg gacctgttcg tccatgctat attggatgtg gatgcagaga atcaaatatt 300
aaagacaagt gtatggtacc aagaggtctg gaatgatgaa tttttatcct ggaactccag 360
catgtttgat gagattagag agatctccct acctctaagt gccatctggg cccccgatat 420
catcatcaat gagtttgtgg acattgaaag ataccctgac cttccctatg tttatgtgaa 480
ctcatctggg accattgaga actataagcc catccaggtg gtctctgcgt gcagtttaga 540
gacatatgct tttccatttg atgtccagaa ttgcagcctg accttcaaga gcattctgca 600
tacagtggaa gacgtagacc tggcctttct gaggagccca gaagacattc agcatgacaa 660
aaaggcgttt ttgaatgaca gtgagtggga acttctatct gtgtcctcca catacagcat 720
cctgcagagc agcgctggag gatttgcaca gattcagttt aatggcactt cttcaccatc 780
tgcatggcct tcttggttct cagcttagct aagtccatcg tgttggtcaa attcctccat 840
gatgagcagc gtggtggaca ggagcagccc ttcttgtgcc ttcgagggga caccgatgct 900
gacaggccta gagtggaacc cagggcccaa cgtgctgtgg taacagagtc ctcgctgtat 960
ggagagcacc tggcccagcc aggaaccctg aaggaagtct ggtcgcagct tcaatctatc 1020
agcaactacc tccaaactca ggaccagaca gaccaacagg aggcagagtg gctggtcctc 1080
ctgtcccgct ttgaccgact gctcttccaa agctaccttt tcatgctggg gatctacacc 1140
atcactctgt gctccctctg ggcactgtgg ggcggcgtgt gaagactgaa gtgttcttca 1200
gtaattgtgc tggcacttag gagagagagg agggggaata atagtgggtt aaaaagcttt 1260
ctgggtcggg tgtggtggtt cttgcctat 1289

<210> 73
<211> 1358
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7504530CB1

<400> 73
cgctcgaggc cgccctggga gagctgagag gcaggctctg gactggggac acagggatag 60
ctgagcccca gctgggggtg gaagctgagc cagggacagt cacggaggaa caagatcaag 120
atgcgctgta actgagaagc ccccaaggcg gaggctgaga atcagagaca tttcagcaga 180
catctacaaa tctgaaagac aaaacatggt tcaagcatcc gggcacaggc ggtccacccg 240

tggctccaaa atggtctcct ggtccgtgat agcaaagatc caggaaatac tgcagaggaa 300
gatggtgcga gagttcctgg ccgagttcat gagcacatat gtcatgatgg tattcggcct 360
tggttccgtg gcccatatgg ttctaaataa aaaatatggg agctaccttg gtgtcaactt 420
gggttttggc ttcggagtca ccatgggagt gcacgtggca ggccgcatct ctggagccca 480
catgaacgca gctgtgacct ttgctaactg tgcgctgggc cgcgtgccct ggaggaagtt 540
tccggtctat gtgctggggc agttcctggg ctccttcctg gcggctgcca ccatctacag 600
tctcttctac acggccattc tccacttttc gggtggacag ctgatggtga ccggtcccgt 660
cgctacagct ggcatttttg ccacctacct tcctgatcac atgacattgt ggcggggctt 720
cctgaatgag gcgtggctga ccgggatgct ccagctgtgt ctcttcgcca tcacggacca 780
ggagaacaac ccagcactgc caggaacaga ggcgctggtg ataggcatcc tcgtggtcat 840
catcggggtg tcccttggca tgaacacagg atatgccatc aacccgtccc gggacctgcc 900
cccccgcatc ttcaccttca ttgctggttg ggcaaacag gtcttcaggt ggcatcatct 960
acctggtctt cattggctcc accatcccac gggagcccct gaaattggag gattctgtgg 1020
cgtatgaaga ccacgggata accgtattgc ccaagatggg atctcatgaa cccacgatct 1080
ctcccctcac ccccgtctct gtgagccctg ccaacagatc ttcagtccac cctgccccac 1140
ccttacatga atccatggcc ctagagcact tctaagcaga gattatttgt gatcccatcc 1200
attccccaat aaagcaaggc ttgtccgaca gcagtacccc cacttcctgg gggcctcctg 1260
tggttgggct tccctcctgg gttcttacag gagctccagg gctatgtctt agcccaaggg 1320
gtagaggtga ggcacctaag gccttccatc cccgggag 1358

<210> 74
<211> 2232
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509303CB1

<400> 74
ggaccctaga cctctgcagc ccataccagg tctcatggag gggaacaagc tggaggagca 60
ggactctagc cctccacagt ccactccagg gctcatgaag gggaacaagc gtgaggagca 120
ggggctgggc cccgaacctg cggcgcccca gcagcccacg gcggaggagg aggccctgat 180
cgagttccac cgctcctacc gagagctctt cgagttcttc tgcaacaaca ccaccatcca 240
cggcgccatc cgcctggtgt gctcccagca caaccgcatg aagacggcct tctgggcagt 300
gctgtggctc tgcacctttg gcatgatgta ctggcaattc ggcctgcttt tcggagagta 360
cttcagctac cccgtcagcc tcaacatcaa cctcaactcg gacaagctcg tcttccccgc 420
agtgaccatc tgcaccctca atccctacag gtacccggaa attaaagagg agctggagga 480
gctggaccgc atcacagagc agacgctctt tgacctgtac aaatacagct ccttcaccac 540
tctcgtggcc ggctcccgca gccgtcgcga cctgcggggg actctgccgc accccttgca 600
gcgcctgagg gtcccgcccc cgcctcacgg ggcccgtcga gcccgtagcg tggcctccag 660
cttgcgggac aacaaccccc aggtggactg gaaggactgg aagatcggct tccagctgga 720
attactctca cttccaccac ccgatgtatg gaaactgcta tactttcaat gacaagaaca 780
actccaacct ctggatgtct tccatgcctg gaatcaacaa cggtctgtcc ctgatgctgc 840
gcgcagagca gaatgacttc attcccctgc tgtccacagt gactggggcc cgggtaatgg 900
tgcacgggca ggatgaacct gcctttatgg atgatggtgg ctttaacttg cggcctggcg 960
tggagacctc catcagcatg aggaaggaaa ccctggacag acttgggggc gattatggcg 1020
actgcaccaa gaatggcagt gatgttcctg ttgagaacct ttacccttca aagtacacac 1080
agcaggtgtg tattcactcc tgcttccagg agagcatgat caaggagtgt ggctgtgcct 1140
acatcttcta tccgcggccc cagaacgtgg agtactgtga ctacagaaag cacagttcct 1200
gggggtactg ctactataag ctccaggttg acttctcctc agaccacctg ggctgtttca 1260
ccaagtgccg gaagccatgc agcgtgacca gctaccagct ctctgctggt tactcacgat 1320
ggccctcggt gacatcccag gaatgggtct tccagatgct atcgcgacag aacaattaca 1380
ccgtcaacaa caagagaaat ggagtggcca aagtcaacat cttcttcaag gagctgaact 1440
acaaaaccaa ttctgagtct ccctctgtca cgatggtcac cctcctgtcc aacctgggca 1500
gccagtggag cctgtggttc ggctcctcgg tgttgtctgt ggtggagatg gctgagctcg 1560

tctttgacct gctggtcatc atgttcctca tgctgctccg aaggttccga agccgatact 1620
ggtctccagg ccgagggggc aggggtgctc aggaggtagc ctccaccctg gcatcctccc 1680
ctccttccca cttctgcccc caccccatgt ctctgtcctt gtcccagcca ggccctgctc 1740
cctctccagc cttgacagcc cctccccctg cctatgccac cctgggcccc cgcccatctc 1800
caggggggctc tgcaggggcc agttcctcca cctgtcctct ggggggggccc tgagagggaa 1860
ggagaggttt ctcacaccaa ggcagatgct cctctggtgg gagggtgctg gccctggcaa 1920
gattgaagga tgtgcagggc ttcctctcag agccgcccaa actgccgttg atgtgtggag 1980
gggaagcaag atgggtaagg gctcaggaag ttgctccaag aacagtagct gatgaagctg 2040
cccagaagtg ccttggctcc agccctgtac cccttggtac tgcctctgaa cactctggtt 2100
tccccaccca actgcggcta agtctctttt tcccttggat cagccaagcg aaacttggag 2160
ctttgacaag gaactttcct aagaaaccgc tgataaccag gacaaaacac aaccaaggta 2220
cacgcaggca tg 2232

<210> 75
<211> 2230
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509910CB1

<400> 75
gttactggga gggggcttgc tgtggccctg tcaggaagag tagagctctg gtccagctcc 60
gcgcagggag ggaggctgtc accatgccgg cctgctgcag ctgcagtgat gttttccagt 120
atgagacgaa caaagtcact cggatccaga gcatgaatta tggcaccatt aagtggttct 180
tccacgtgat catcttttcc tacgtttgct ttgctctggt gagtgacaag ctgtaccagc 240
ggaaagagcc tgtcatcagt tctgtgcaca ccaaggtgaa ggggatagca gaggtgaaag 300
aggagatcgt ggagaatgga gtgaagaagt tggtgcacag tgtctttgac accgcagact 360
acaccttccc tttgcagggg aactctttct tcgtgatgac aaactttctc aaaacagaag 420
gccaagagca gcggttgtgt cccgagtatc ccacccgcag gacgctctgt tcctctgacc 480
gaggttgtaa aaagggatgg atggacccgc agagcaaagg aattcagacc ggaaggtgtg 540
tagtgcatga agggaaccag aagacctgtg aagtctctgc ctggtgcccc atcgaggcag 600
tggaagaggc ccccccggcct gctctcttga acagtgccga aaacttcact gtgctcatca 660
agaacaatat cgacttcccc ggccacaact acaccacgag aaacatcctg ccaggtttaa 720
acatcacttg taccttccac aagactcaga atccacagtg tcccattttc cgactaggag 780
acatcttccg agaaacaggc gataattttt cagatgtggc aattcagggc ggaataatgg 840
gcattgagat ctactgggac tgcaacctag accgttggtt ccatcactgc cgtcccaaat 900
acagtttccg tcgccttgac gacaagacca ccaacgtgtc cttgtaccct ggctacaact 960
tcagatacgc caagtactac aaggaaaaca atgttgagaa acggactctg ataaaagtct 1020
tcgggatccg ttttgacatc ctggtttttg gcaccggagg aaaatttgac attatccagc 1080
tggttgtgta catcggctca accctctcct acttcggtct ggtaagagat tctcttttcc 1140
atgctttagg aaaatggttt ggagaaggaa gtgactaacg cagcgcttgt ctgcattctc 1200
cccaggccgc tgtgttcatc gacttcctca tcgacactta ctccagtaac tgctgtcgct 1260
cccatattta tccctggtgc aagtgctgtc agccctgtgt ggtcaacgaa tactactaca 1320
ggaagaagtg cgagtccatt gtggagccaa agccgacatt aaagtatgtg tcctttgtgg 1380
atgaatccca cattaggatg gtgaaccagc agctactagg gagaagtctg caagatgtca 1440
agggccaaga agtcccaaga cctgcgatgg acttcacaga tttgtccagg ctgcccctgg 1500
ccctccatga cacaccccccg attcctggac aaccagagga gatacagctg cttagaaagg 1560
aggcgactcc tagatccagg gatagccccg tctggtgcca gtgtggaagc tgcctcccat 1620
ctcaactccc tgagagccac aggtgcctgg aggagctgtg ctgccggaaa aagccggggg 1680
cctgcatcac cacctcagag ctgttcagga agctggtcct gtccagacac gtcctgcagt 1740
tcctcctgct ctaccaggag cccttgctgg cgctggatgt ggattccacc aacagccggc 1800
tgcggcactg tgcctacagg tgctacgcca cctggcgctt cggctcccag gacatggctg 1860
actttgccat cctgcccagc tgctgccgct ggaggatccg gaaagagttt ccgaagagtg 1920
aagggcagta cagtggcttc aagagtcctt actgaagcca ggcaccgtgg ctcacgtctg 1980

taatcccagc gctttgggag gccgaggcag gcagatcacc tgaggtcggg agttggagac 2040
ccgcctggct aacaaggcga aatcctgtct gtactaaaaa tacaaaaatc agccagacat 2100
ggtggcatgc acctgcaatc ccagctactc gggaggctga ggcacaagaa tcacttgaac 2160
ccgggaggca gaggttgtag tgagcccaga ttgtgccact gctctccagc ctgggaggca 2220
cagcaaactg 2230

<210> 76
<211> 5966
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509982CB1

<400> 76
gcttggtctg gctacacccc ttttcagaaa ccaggctgtg taagagctgc tggagtaggc 60
acccatttaa agaaaaaatg aagaagcagc aataaagaag ttgtaatcgt tacctagaca 120
aacagagaac tggttttgac agtgtttcta gagtgctttt tattattttc ctgacagttg 180
tgttccacca tgattacttt ctccttcagc gaataggcta aatgaatatg aaacagaaaa 240
gcgtgtatca gcaaaccaaa gcacttctgt gcaagaattt tcttaagaaa tggaggatga 300
aaagagagag cttattggaa tggggcctct caatacttct aggactgtgt attgctctgt 360
tttccagttc catgagaaat gtccagtttc ctggaatggc tcctcagaat ctgggaaggg 420
tagataaatt taatagctct tctttaatgg ttgtgtatac accaatatct aatttaaccc 480
agcagataat gaataaaaca gcacttgctc ctcttttgaa aggaacaagt gtcattgggg 540
caccaaataa aacacacatg gacgaaatac ttctggaaaa tttaccatat gctatgggaa 600
tcatctttaa tgaaactttc tcttataagt taatattttt ccagggatat aacagtccac 660
tttggaaaga agatttctca gctcattgct gggatggata tggtgagttt tcatgtacat 720
tgaccaaata ctggaataga ggatttgtgg ctttacaaac agctattaat actgccatta 780
tagaaatcac aaccaatcac cctgtgatgg aggagttgat gtcagttact gctataacta 840
tgaagacatt acctttcata actaaaaatc ttcttcacaa tgagatgttt attttattct 900
tcttgcttca tttctcccca cttgtatatt ttatatcact caatgtaaca aaagagagaa 960
aaaagtctaa gaatttgatg aaaatgatgg gtctccaaga ttcagcattc tggctctcct 1020
ggggtctaat ctatgctggc ttcatcttta ttatttccat attcattaca attatcataa 1080
cattcaccca aattatagtc atgactggct tcatggtcat atttataccc ttttttttat 1140
atggcttatc tttggtagct ttggtgttcc tgctgagtgt gctgttaaag aaagctgtcc 1200
tcaccaattt ggttgtgttt ctccttaccc tcttttgggg atgtctggga ttcactgtat 1260
tttatgaaca acttccttca tctctggagt ggattttgaa tatttgtagc ccttttgcct 1320
ttactactgg aatgattcag attatcaaac tggattataa cttgaatggt gtaatttttc 1380
ctgacccttc aggagactca tatacaatga tagcaacttt ttctatgttg cttttggatg 1440
gtctcatcta cttgctattg gcattatact ttgacaaaat tttaccctat ggagatgagc 1500
gccattattc tcctttattt ttcttgaatt catcatcttg tttccaacac caaaggacta 1560
atgctaaggt tattgagaaa gaaatcgatg ctgagcatcc ctctgatgat tattttgaac 1620
cagtagctcc tgaattccaa ggaaaagaag ccatcagaat cagaaatgtt aagaaggaat 1680
ataaaggaaa atctggaaaa gtggaagcat tgaaaggctt gctctttgac atatatgaag 1740
gtcaaatcac ggcaatcctg ggtcacagtg gagctggcaa atcttcactg ctaaatattc 1800
ttaatggatt gtctgttcca acagaaggat cagttaccat ctataataaa aatctctctg 1860
aaatgcaaga cttggaggaa atcagaaaga taactggcgt ctgtcctcaa ttcaatgttc 1920
aatttgacat actcaccgtg aaggaaaacc tcagcctgtt tgctaaaata aaagggattc 1980
atctaaagga agtggaacaa gaggtacaac gaatattatt ggaattggac atgcaaaaca 2040
ttcaagataa ccttgctaaa catttaagtg aaggacagaa aagaaagctg acttttggga 2100
ttaccatttt aggagatcct caaattttgc ttttagatga accaactact ggattggatc 2160
ccttttccag agatcaagtg tggagcctcc tgagagagcg tagagcagat catgtgatcc 2220
ttttcagtac ccagtccatg gatgaggctg acatcctggc tgatagaaaa gtgatcatgt 2280
ccaatgggag actgaagtgt gcaggttctt ctatgttttt gaaaagaagg tggggtcttg 2340
gatatcacct aagtttacat aggaatgaaa tatgtaaccc agaacaaata acatccttca 2400

ttactcatca catccccgat gctaaattaa aaacagaaaa caaagaaaag cttgtatata 2460
ctttgccact ggaaaggaca aatacatttc cagatctttt cagtgatctg gataagtgtt 2520
ctgaccaggg agtgacaggt tatgacattt ccatgtcaac tctaaatgaa gtctttatga 2580
aactggaagg acagtcaact atcgaacaag gtaaagccat ttgtataaat ttcgaacaag 2640
tggagatgat aagagactca gaaagcctca atgaaatgga gctggctcac tcttccttct 2700
ctgaaatgca gacagctgtg agtgacatgg gcctctggag aatgcaagtc tttgccatgg 2760
cacggctccg tttcttaaag ttaaaacgtc aaactaaagt gttattgacc ctattattgg 2820
tatttggaat cgcaatattc cctttgattg ttgaaaatat aatatatgct atgttaaatg 2880
aaaagatcga ttgggaattt aaaaacgaat tgtattttct ctctcctgga caacttcccc 2940
aggaaccccg taccagcctg ttgatcatca ataacacaga atcaaatatt gaagatttta 3000
taaaatcact gaagcatcaa aatatacttt tggaagtaga tgactttgaa aacagaaatg 3060
gtactgatgg cctctcatac aatggagcta tcatagtttc tggtaaacaa aaggattata 3120
gattttcagt tgtgtgtaat accaagagat tgcactgttt tccaattctt atgaatatta 3180
tcagtaatgg gctacttcaa atgtttaatc acacacaaca tattcgaatt gagtcaagcc 3240
catttcctct tagccacata ggactctgga ctgggttgcc ggatggttcc tttttcttat 3300
ttttggttct atgtagcatt tctccttata tcaccatggg cagcatcagt gattacaaga 3360
aaaatgctaa gtcccagcta tggatttcag gcctctacac ttctgcttac tggtgtgggc 3420
aggcactagt ggacgtcagc ttcttcattt taattctcct tttaatgtat ttaattttct 3480
acatagaaaa catgcagtac cttcttatta caagccaaat tgtgtttgct ttggttatag 3540
ttactcctgg ttatgcagct tctcttgtct tcttcatata tatgatatca tttatttttc 3600
gcaaaaggag aaaaaacagt ggcctttggt cattttactt ctttttttgcc tccaccatca 3660
tgttttccat cactttaatc aatcattttg acctaagtat attgattacc accatggtat 3720
tggttccttc atataccttg cttggattta aacttttttt ggaagtgaga gaccaggagc 3780
actacagaga atttccagag gcaaattttg aattgagtgc cactgatttt ctagtctgct 3840
tcataccctа ctttcagact ttgctattcg tttttgttct aagatacatg gaactaaaat 3900
gtggaaagaa aagaatgcga aaagatcctg ttttcagaat ttcccccсаа agtagagatg 3960
ctaagccaaa tccagaagaa cccatagatg aagatgaaga tattcaaaca gaaagaataa 4020
gaacagtcac tgctctgacc acttcaatct tagatgagaa acctgttata attgccagct 4080
gtctacacaa agaatatgca ggccagaaga aaagttgctt ttcaaagagg aagaagaaaa 4140
tagcagcaag aaatatctct ttctgtgttc aagaaggtga gattttggga ttgctaggac 4200
ccagtggtgc tggaaaaagt tcatctatta gaatgatatc tgggatcaca aagccaactg 4260
ctggagaggt ggaactgaaa ggctgcagtt cagttttggg ccacctgggg tactgccctc 4320
aagagaacgt gctgtggccc atgctgacgt tgagggaaca cctggaggtg tatgctgccg 4380
tcaaggggct cagggaagcg gacgcgaggc tcgccatcgc aagattagtg agtgctttca 4440
aactgcatga gcagctgaat gttcctgtgc agaaattaac agcaggaatc acgagaaagt 4500
tgtgttttgt gctgagcctc ctgggaaact cacctgtctt gctcctggat gaaccatcta 4560
cgggcataga ccccacaggg cagcagcaaa tgtggcaggc aatccaggca gtcgttaaaa 4620
acacagagag aggtgtcctc ctgaccaccc ataacctggc tgaggcggaa gccttgtgtg 4680
accgtgtggc catcatggtg tctggaaggc ttagatgcat tggctccatc caacacctga 4740
aaaacaaact tggcaaggat tacattctag agctaaaagt gaaggaaacg tctcaagtga 4800
ctttggtcca cactgagatt ctgaagcttt tcccacaggc tgcagggcag gaaaggtatt 4860
cctctttgtt aacctataag ctgcccgtgg cagacgttta ccctctatca cagacctttc 4920
acaaattaga agcagtgaag cataacttta acctggaaga atacagcctt tctcagtgca 4980
cactggagaa ggtattctta gagctttcta aagaacagga agtaggaaat tttgatgaag 5040
aaattgatac aacaatgaga tggaaactcc tccctcattc agatgaacct taaaacctca 5100
aacctagtaa tttttttgttg atctcctata aacttatgtt ttatgtaata attaatagta 5160
tgtttaattt taaagatcat ttaaaattaa catcaggtat attttgtaaa tttagttaac 5220
aaatacataa attttaaaat tattcttcct ctcaaacata ggggtgatag caaacctgtg 5280
ataaaggcaa tacaaaatat tagtaaagtc acccaaagag tcaggcactg ggtattgtgg 5340
aaataaaact atataaactt agaatttttt aaaaatatga ctttttttacc ttttacaaaa 5400
cattctcttg ctgaaatatg tgaagggtat attcagtagc caagagttgc atgactactt 5460
cacaccagtt catgatacaa caggtataca ggttttcttt tataaccaac tacaactcaa 5520
gagtcttctg aaagtgttcc agaaattgct ttaaaactca aaagtaaggg gccaggtgca 5580
gtggctcacg cctgtaatcc cagcactttg ggaggccgag gcaggtggat cacaaggtca 5640
ggagttcgag actagcctgg ccaatatggt gaaactccat ctctaataaa aatacaaaaa 5700
ttagccgggc gttggcattt gcctgtagtc ctagctattc gggaggctga gggaggagaa 5760

ttgcttgaac ccgggaggca gaggttgcag tgagccatgt gctagtgcac tccagcctgg 5820
gtgacagagt gagactctgt caaaaaaaaa aaaaacaaaa aaaaaaacaa aaaaccttca 5880
aggttttgga ggtctttggc cacaatttga gagccccttt tggaaaggtt tccctttttac 5940
ttttgaataa agggtccgga tttggc 5966

<210> 77
<211> 2071
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510082CB1

<400> 77
ccgggtagtg agcggaggga caggaagggt agggcaagaa agggagaggg gacaggaggg 60
aagggtgggc caaagcggtg agaaaggagg gccagccagt tgcgtggggg agagtggccg 120
aggcccgggg gcaggagtgc agggctctga ggcggggaga ggagaggaga gaagagccgc 180
ggggggccca gcccggagcc aggatgcccg cgccgcgcgc ccgggagcag ccccgcgtgc 240
ccggggagcg ccagccgctg ctgcctcgcg gtgcgcgggg ccctcgacgg tggcggcggg 300
cggcgggcgc ggccgtgctg ctggtggaga tgctggagcg cgccgccttc ttcggcgtca 360
ccgccaacct cgtgctgtac ctcaacagca ccaacttcaa ctggaccggc gagcaggcga 420
cgcgcgccgc gctggtattc ctgggcgcct cctacctgct ggcgcccgtg ggcggctggc 480
tggccgacgt gtacctgggc cgctaccgcg cggtcgcgct cagcctgctg ctctacctgg 540
ccgcctcggg cctgctgccc gccaccgcct tccccgacgg ccgcagctcc ttctgcggag 600
agatgcccgc gtcgccgctg ggacctgcct gcccctcggc cggctgcccg cgctcctcgc 660
ccagccccta ctgcgcgccc gtcctctacg cgggcctgct gctactcggc ctggccgcca 720
gctccgtccg gagcaacctc acctccttcg gtgccgacca ggtgatggat ctcggccgcg 780
acgccacccg ccgcttcttc aactggtttt actggagcat caacctgggt gctgtgctgt 840
cgctgctggt ggtggcgttt attcagcaga acatcagctt cctgctgggc tacagcatcc 900
ctgtgggctg tgtgggcctg gcattttca tcttcctctt tgccaccccc gtcttcatca 960
ccaagccccc gatgggcagc caagtgtcct ctatgcttaa gctcgctctc caaaactgct 1020
gcccccagct gtggcaacga cactcggcca gagaccgtca atgtgcccgc gtgctggccg 1080
acgagaggtc tccccagcca ggggcttccc cgcaagagga catcgccaac ttccaggtgc 1140
tggtgaagat cttgcccgtc atggtgaccc tggtgcccta ctggatggtc tacttccaga 1200
tgcagtccac ctatgtcctg cagggtcttc acctccacat cccaaacatt ttcccagcca 1260
acccggccaa catctctgtg gccctgagag cccagggcag cagctacacg atcccggaag 1320
cctggctcct cctggccaat gttgtggtgg tgctgattct ggtccctctg aaggaccgct 1380
tgatcgaccc tttactgctg cggtgcaagc tgcttccctc tgctctgcag aagatggcgc 1440
tggggatgtt ctttggtttt acctccgtca ttgtggcagg agtcctggag atggagcgct 1500
tacactacat ccaccacaac gagaccgtgt cccagcagat tggggaggtc ctgtacaacg 1560
cggcaccact gtccatctgg tggcagatcc ctcagtacct gctcattggg atcagtgaga 1620
tctttgccag catcccaggc ctggagtttg cctactcaga ggccccgcgc tccatgcagg 1680
gcgccatcat gggcatcttc ttctgcctgt cgggggtggg ctcactgttg ggctccagcc 1740
tagtggcact gctgtccttg cccgggggct ggctgcactg ccccaaggac tttgggaaca 1800
tcaacaattg ccggatggac ctctacttct tcctgctggc tggcattcag gccgtcacgg 1860
ctctcctatt tgtctggatc gctggacgct atgagagggc gtcccagggc ccagcctccc 1920
acaggccgtt tcagcacgga caggggctcg acaggcccta tccaggcccc ttggtttact 1980
ctaccggaaa gaacggcagc agtcccagct ctggtttcct tctcggttta ttctgttaga 2040
atgaaatgtt cccataaata agggcatggt c 2071

<210> 78
<211> 3703
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510367CB1

<220>
<221> unsure
<222> (1) ... (3703)
<223> a, t, c, g, or other

<400> 78
cggacggtgg gcggcgtgac gtagtggctg tgccccgttc ttgccccctc agtactagag 60
tctccggctt cgctcacgcg ccttgggcat aagagtcctc tcgttggtcc cggaggtggg 120
gttgcgctca caaggggcga ccgtcgccac ggtggcggcc actgcatcgc gtcccacctc 180
cgcggccctg ggcgccgtgg tgtcgacggg ccccgagcct atgacgggcc agggccagtc 240
ggcgtccggg tcgtcggcgt ggagcacggt attccgccac gtccggtatg agaacctgat 300
agcgggcgtg agcggcggcg tcttatccaa ccttgcgctg catccgctcg acctcgtgaa 360
gatccgcttc gccggaacaa tcctgtgagt ttatgatatc tgctatacag atgaggataa 420
cagaagttga ctaacttgcc caagtcacaa gtttgtaaag atggatctgg gatttaaacc 480
cagtgagtga tggattggaa ctgagaccga aatataatgg aattttacat tgcttgacta 540
ccatttggaa acttgatgga ctacggggac tttatcaagg agtaacccca aatatatggg 600
gtgcaggttt atcctgggga ctctactttt tcttttacaa tgccatcaag tcatataaaa 660
cagaaggaag agctgaacgt ttagaggcaa cagaatacct tgtctcagct gctgaagctg 720
gagccatgac cctctgcatt acaaacccat tatgggtaac aaaaactcgc cttatgttac 780
agtatgatgc tgttgttaac tccccacacc gacaatataa aggaatgttt gatacacttg 840
tgaaaatata taagtatgaa ggtgtgcgtg gattatataa gggatttgtt cctgggctgt 900
ttggaacatc gcatggtgcc cttcagttta tggcatatga attgctgaag ttgaagtaca 960
accagcatat caatagatta ccagaagccc agttgagcac agtagaatat atatctgttg 1020
cagcactatc caaaatattt gctgtcgcag caacataccc atatcaagtc gtaagagctc 1080
gtcttcagga tcaacacatg ttttacagtg gtgtaataga tgtaatcaca aagacatgga 1140
ggaaagaagg cgtcggtgga ttttacaagg gaattgctcc taatttgatt agagtgactc 1200
cagcctgctg tattaccttt gtggtatatg aaaacgtctc acattttta cttgacctta 1260
gagaaaagag aaagtaagct caaagaggac aattccagta tatctgccca aggcagcaac 1320
aagctctttt gtgtttaagg cataaaagaa gaattctgca tagaaacatg gctcatattc 1380
gaaattgctc tatagtcatt agaagccaga gaactgctaa gtctcctgca atgtttttct 1440
gctttttgcc ttccccatat atatggaact tggctacctc tgcctgaaat ggctgccatc 1500
aacacaatgt taaaactgac acgaaggata gagtttcaca gatttctacg ttttattggt 1560
ggaagctgat ttgcaacatt tgctaaatgg attagatgaa tgtacttctt tttgtgagct 1620
tacttgcctg gattgcttta aaattaacct ttgtgcaata ccaagaaaat agctctttaa 1680
aagaatgtct ttgtatgtct caaggtaaat taaggattta ctgaataagg tgttgaccaa 1740
atccagacca ttttatttta tttttttatt tatttatttt ttgagatgga gtcttgcttt 1800
gtcgcccagg ctggagtgca gtggcgtgat ctcagctcac tgcaacctcc acctcccggg 1860
ttcacgccat tctcctgcct cagcctcctg agtagctggg actacaggca cctgccacca 1920
cgcctggcta acttttttt atattttgag tagaaatggg gtttcaccat gttagccagg 1980
atggtctcaa tctcctgacc ttgtgatccg cctgccttgg cctcccaaag tgctgggatt 2040
acaggcgtga gccactgcgc ctggccagac cattttagaa ttgggaaatt ttagtgagaa 2100
aaaatgcact gtaaatatgc tttagtttta attcagttgg gatgcactac ctagcgaaaa 2160
ttgagaaact atatacttct cagagaaata tctgacatct attgtcattc cattgctatt 2220
ttttttcccc agagacttcc ataatttaaa ataaaatcct agatccagtt cttgtttttt 2280
ggcataaata cttaatctat tttaaattta taaaatctga gcttctagga tccagctgtg 2340
tcaaccttta tttagcatat ataactataa atcacttatt acagatgcta aatagatcac 2400
cttttacaga tgctgaaatg tttgggatat gtttgttgac aaggtaaatg gaaatgagaa 2460
actttatact tcagttttca gatatatgga tctagatccc aaataaatga ttaatcttca 2520
ttggtttctc aaattcaggt tgaaatacaa attaatagcc tttattgatt ttacttttat 2580
gagtcattgt agacatctat aaatataaaa gggcctgtac ccaaaggatg ccggaatact 2640
agtattttta tttatcgtaa acatccacga gtgctgttgc actaccatct atttgttgta 2700
aataaaagtg ttgttttcaa agccatcttt aaatagttct ttaaaaatag gtcttttttt 2760

tatattttgg aaaaggcatt gtttttaaag taaagataaa atggtaagta cctaattgta 2820
tttactgtaa tataacatgc agatgaatgc tttataagtt aaatatgatg tatttttca 2880
tacttctgga ttatactata attcatatga aatcttgata ttagtcccca cacggaaaaa 2940
gtgaactgca gttgatattt ggtgtttaag atagcaccat tgtttaaata ccgcctatgt 3000
actcccaaat gaataaaaca taattcttgt cctctgagag catacaagct tgtgtgatag 3060
ataaatatgc attaaataat tacagctata gattaagaac tcgtaaggaa tatctcacaa 3120
agcctggtaa gagttctgac agaagagaac tcaatttcag tcattcaaca acaaatagtc 3180
actcaactcc tgctctttgg caggtacgtc accaagcact gaggatatag caataaataa 3240
gccagacaat attcctgccc tcagggaatt ttcatcttgg tgtaagtgac agaaaaggaa 3300
atatatgcat aatgtaattt caggtattta atgctataaa taaaactaca gcagaacggg 3360
ggaatagtgt tgctatctta ggtagtatgg ttggggtcag gaaaggtctc tgagatagta 3420
tttggaacat tttgaaagaa gtgaggaagc gagccacagt gagaggaaac acattctagg 3480
cagagaggac aggaagtaaa aagggcctga tggaaggact aaagttggat agaggggatg 3540
ggtgagcaca aaacggagtt ggcaaattgg ccttgccgta tgtgtggaaa gacccaggtg 3600
ggacgccang cccctggttc aaaataccgt tcccattctc gtttgccttg gcggatatta 3660
caccctatgt tatggttccg ccttcaaagg cgttaacacc cct 3703

<210> 79
<211> 1171
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510413CB1

<400> 79
gagcacggag attcggcacg aggaggaaaa gcagcaaacc agatggagaa ggaaggccaa 60
aaagataagt agaaagctgg accatgagct ccttgagggc agggactgaa tggttactat 120
gtccccaggg cccagcatga ccttctcctg gattcctcat cttccttctg tgacctgtgt 180
ctccatcagt ttctcctccg gcatcttttt cttacaggat tcttacctca ggaatagatg 240
gacatggcct ggcagatgat gcagctgctg cttctggctt tggtgactgc tgcggggagt 300
gcccagccca ggagtgcgcg ggccaggacg gacctgctca atgtctgcat gaacgccaag 360
caccacaaga cacagcccag ccccgaggac gagctgtatg gccaggtggg agctcctcaa 420
gggccctccc caggaagtgt tcctctggat gacctacctg ggcagagga gccagaatat 480
ggaggagatg gctgtggtgg ggagagactt agtcctgtgt cttccccacc cagtgcagtc 540
cctggaagaa gaatgcctgc tgcacggcca gcaccagcca ggagctgcac aaggacacct 600
cccgcctgta caactttaac tgggatcact gtggtaagat ggaacccacc tgcaagcgcc 660
actttatcca ggacagctgt ctctgagtgc tcacccaacc tggggccctg gatccggcag 720
gtcaaccaga gctggcgcaa agagcgcatt ctgaacgtgc ccctgtgcaa agaggactgt 780
gagcgctggt gggaggactg tcgcacctcc tacacctgca aaagcaactg gcacaaaggc 840
tggaattgga cctcagggat taatgagtgt ccggccgggg ccctctgcag cacctttgag 900
tcctacttcc ccactccagc cgccctttgt gaaggcctct ggagccactc cttcaaggtc 960
agcaactata gtcgagggag cggccgctgc atccagatgt ggtttgactc agcccagggc 1020
aaccccaatg aggaggtggc caagttctat gctgcggcca tgaatgctgg ggccccgtct 1080
cgtgggatta ttgattcctg atccaagaag ggtcctctgg ggttcttcca acaacctatt 1140
ctaatagaca aatccacatg aaaaaaaaaa a 1171

<210> 80
<211> 323
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1721303CB1

<400> 80
ccaagactcc aaaatggcgt cagttggtga gtgtccggcc ccagtaccag tgaaggacaa 60
gaaacttctg gaggtcaaac tgggggagct gccaagctgg atcttgatgc gggacttcag 120
tcctagtggc attttcggag cgtttcaaag agagcacgag cggctccgca aataccactg 180
aagaggacac actctgcacc cccccacccc acgaccttgg cccgagcccc tccgtgagga 240
acacaatctc aatcgttgct gaatcctttc atatcctaat aggaattaac ctccaaataa 300
aacatgactg gtacgtgtaa aaa 323

<210> 81
<211> 1221
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7502007CB1

<400> 81
attaactcag ctgggagttg aagagccgat gggcagcagg cagacttgag tctcctttct 60
gtccatgggc tcgggccact gtctcaggtc cacccgtggc tccaaaatgg tctcctggtc 120
cgtgatagca aagatccagg aaatactgca gaggaagatg gtgcgagagt tcctggccga 180
gttcatgagc acatatgtca tgatggtatt cggccttggt tccgtggccc atatggttct 240
aaataaaaaa tatgggagct accttggtgt caacttgggt tttggcttcg gagtcaccat 300
gggagtgcac gtggcaggcc gcatctctgg agcccacatg aacgcagctg tgacctttgc 360
taactgtgcg ctgggccgcg tgccctggag gaagtttccg gtctatgtgc tggggcagtt 420
cctgggctcc ttcctggcgg ctgccaccat ctacagtctc ttctacacgg ccattctcca 480
cttttcgggt ggacagctga tggtgaccgg tcccgtcgct acagctggca tttttgccac 540
ctaccttcct gatcacatga cattgtggcg gggcttcctg aatgaggcgt ggctgaccgg 600
gatgctccag ctgtgtctct tcgccatcac ggaccaggag aacaacccag cactgccagg 660
aacagaggcg ctggtgatag gcatcctcgt ggtcatcatc ggggtgtccc ttggcatgaa 720
cacaggatat gccatcaacc cgtcccggga cctgcccccc cgcatcttca ccttcattgc 780
tggttggggc aaacaggtct tcaggtggca tcatctacct ggtcttcatt ggctccacca 840
tcccacggga gcccctgaaa ttggaggatt ctgtggcgta tgaagaccac gggataaccg 900
tattgcccaa gatgggatct catgaaccca cgatctctcc cctcaccccc gtctctgtga 960
gccctgccaa cagatcttca gtccaccctg ccccacccct acatgaatcc atggccctag 1020
agcacttcta agcagagatt atttgtgatc ccatccattc cccaataaag caaggcttgt 1080
ccgacaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1140
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaagg ggggggggcg 1200
cgggggcccc gacaaggggg g 1221

<210> 82
<211> 2008
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506439CB1

<400> 82
gcagcctcag aaggtgtgag cagtggccac gagaggcagg ctggctggga catgaggttg 60
gcagagggca ggcaagctgg cccttggtgg gcctcgtcct gagcactcgg aggcactcct 120
atgcttggaa agctcgctat gctgctgtgg gtccagcagg cgctgctcgc cttgctcctc 180
cccacactcc tggcacaggg agaagccagg aggagccgaa acaccaccag gcccgctctg 240
ctgaggctgt cggattacct tttgaccaac tacaggaagg gtgtgcgccc cgtgagggac 300
tggaggaagc caaccaccgt atccattgac gtcattgtct atgccatcct caacgtggat 360

gagaagaatc aggtgctgac cacctacatc tggtaccggc agtactggac tgatgagttt 420
ctccagtgga accctgagga ctttgacaac atcaccaagt tgtccatccc cacggacagc 480
atctgggtcc cggacattct catcaatgag ttcgtggatg tggggaagtc tccaaatatc 540
ccgtacgtgt atattcggca tcaacatctc tttgtggcgc ttgccagaaa aggtgaaatc 600
cgacaggagt gtcttcatga accagggaga gtgggagttg ctgggggtgc tgccctactt 660
tcgggagttc agcatggaaa gcagtaacta ctatgcagaa atgaagttct atgtggtcat 720
ccgccggcgg cccctcttct atgtggtcag cctgctactg cccagcatct tcctcatggt 780
catggacatc gtgggcttct acctgccccc caacagtggc gagagggtct ctttcaagat 840
tacactcctc ctgggctact cggtcttcct gatcatcgtt tctgacacgc tgccggccac 900
tgccatcggc actcctctca ttggtgtcta ctttgtggtg tgcatggctc tgctggtgat 960
aagtttggcc gagaccatct tcattgtgcg gctggtgcac aagcaagacc tgcagcagcc 1020
cgtgcctgct tggctgcgtc acctggttct ggagagaatc gcctggctac tttgcctgag 1080
ggagcagtca acttcccaga ggcccccagc cacctcccaa gccaccaaga ctgatgactg 1140
ctcagccatg ggaaaccact gcagccacat gggaggaccc caggacttcg agaagagccc 1200
gagggacaga tgtagccctc ccccaccacc tcgggaggcc tcgctggcgg tgtgtgggct 1260
gctgcaggag ctgtcctcca tccggcaatt cctggaaaag cgggatgaga tccgagaggt 1320
ggcccgagac tggctgcgcg tgggctccgt gctggacaag ctgctattcc acatttacct 1380
gctggcggtg ctggcctaca gcatcaccct ggttatgctc tggtccatct ggcagtacgc 1440
ttgagtgggt acagcccagt ggaggagggg gtacagtcct ggttaggtgg ggacagagga 1500
tttctgctta ggcccctcag gacccaggga atgccaggga cattttcaag acacagacaa 1560
agtcccgtgc cctgtttcca atgccaattc atctcagcaa tcacaagcca aggtctgaac 1620
ccttccacca aaaactgggt gttcaaggcc cttacaccct tgtcccaccc ccagcagctc 1680
accatggctt taaaacatgc tctcttagat caggagaaac tcgggcactc cctaagtcca 1740
ctctagttgt ggactttccc ccattgaccc tcacctgaat aagggacttt ggaattctgc 1800
ttctctttca caactttgct tttaggttga aggcaaaacc aactctctac tacacaggcc 1860
tgataactct gtacgaggct tctctaaccc ctagtgtctt ttttttcttc acctcacttg 1920
tggcagcttc cctgaacact catcccccat cagatgatgg gagtgggaag aataaaatgc 1980
agtgaaaccc taaaaaaaaa aaaaaggg 2008

<210> 83
<211> 1080
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509243CB1

<400> 83
ggaggagggc agggcaggcg gcagcctggg cacaggcccc taggtgctta ctcctcacct 60
gtttcccacc tctcccccat agccagcccc acggccctgg cagggtcctg gccacagcat 120
gcccagtgct gggctctgca gctgctgggg tggccgggtg ctgcccctgc tgctggccta 180
tgtctgctac ctgctgctcg gtgccactat cttccagctg ctagagaggc aggcggaggc 240
tcagtccagg gaccagtttc agttggagaa gctgcgcttc ctggagaact acacctgcct 300
ggaccagtgg gccatggagc agtttgtgca ggtcatcatg gaagcctggg tgaaaggtgt 360
gaaccccaaa ggcaactcta ccaaccccag caactgggac tttggcagca gtttcttctt 420
tgcaggcaca gtcgtcacta ccatagggca ctgacagcag ctgctacacc ccctcggcca 480
gaaaactcac tccatcccac cccaaggaaa ccatcagatc caggatatgg gaacctggca 540
cccagcacag aggcaggtta ggtcttctgt gtcttctatg ccctgttggg catcccgctt 600
aacgtgatct tcctcaacca cctgggcaca gggctgcgtg cccatctggc cgccattgaa 660
agatgggagg accgtcccag gcgctcccag gtactgcaag tcctgggcct ggctctgttc 720
ctgaccctgg ggacgctggt cattctcatc ttcccaccca tggtcttcag ccatgtggag 780
ggctggagct tcagcgaggg cttctacttt gctttcatca ctctcagcac cattggcttt 840
ggggactatg ttgttggcac agaccccagc aagcattata tctcagtgta tcggagcctg 900
gcagccatct ggatcctcct gggcctggcg tggctggcgc tgatcctccc actgggcccc 960
ctgcttctgc acagatgctg ccagctctgg ctgctcagta ggggcctcgg cgtcaaggat 1020

ggggcagcct ctgaccccag tgggctcccc aggcctcaga agatccccat ctctgcatga 1080

<210> 84
<211> 2412
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509404CB1

<400> 84
ctcgagccgc tcgagccgcg ggcggctggg acgcagctgc ccccgctcgg cgcccgctgc 60
acccttagca gccactgcca cctgggcccg gagcctccac gatctcgccc ggcgattgtg 120
ggcaggggcg cctccggatc gatcttctga aattcaagtt ttcaagatga agtttttatt 180
gacaactgcc tttttaattt taatttcctt gtgggtggaa gaagcctatt ctaaggaaaa 240
gtcttcaaag aaagggaagg ggaaaaagaa gcagtatcta tgcccatcag agagattgta 300
ccattgaatg acagcatttt catacttcag tctgtatgta ctaagaaggt cagcagtcag 360
cagaggacct tgcccgagta cctgccaact ccactagcaa tatcttgaac aggttattgg 420
tcagttatga tcccaggata agaccaaact tcaaaggcat tcctgttgat gtagtagtca 480
acatttttat taacagtttt ggatccattc aagaaacaac aatggactat agagttaaca 540
tcttcctgag acaaaaatgg aatgacccca ggctgaagct ccccagtgat tttaggggtt 600
cagatgcact gacagtggat ccaacaatgt acaagtgttt atggaaacct gatttatttt 660
ttgcaaatga aaaaagtgcc aattttcatg atgtgaccca ggaaaacatc ctcctcttta 720
tttttcgtga tggagatgtc cttgtcagca tgaggttatc tattactctt tcatgccctt 780
tggacttgac attgtttccc atggatacac aacgttgcaa gatgcaactg gagagctttg 840
gttacacaac tgatgattta cgatttatct ggcagtcagg agatcctgtg caattagaaa 900
aaattgcctt gcctcaattt gatatcaaaa aggaagatat tgaatatggt aactgtacaa 960
aatactataa aggcacgggc tactacacat gcgtggaagt catcttcacc ctgaggaggc 1020
aggtcggctt ttacatgatg gggtctacg ccccaaccct gctcattgtt gttctctcct 1080
ggctttcctt ctggatcaac ccggacgcga gtgctgccag agtgcccctg ggttggtgag 1140
accagatgca aaaaagtttg tacttctaag tctgatctga gatctaatga cttcagcatt 1200
gttggaagct taccaagaga ttttgaacta tccaattatg actgctatgg aaaacccatt 1260
gaagttaaca acggacttgg gaaatctcag gctaagaaca acaagaagcc tccccctgcg 1320
aaacctgtta ttccaacagc agcaaagcga attgatcttt atgcaagagc attgtttcct 1380
ttctgcttct tgttcttcaa tgttatatat tggtctatat atttatgata aatcttttcc 1440
atttgtacaa aataaaattc catttcattg tgacctactc ctttcataaa tgccaatctg 1500
tgagaacttt tgaattttca tagcaacatt gcattttgga tgccatttga ttgtaataaa 1560
actgtggcac cttaattttg aatggcagca tgatcatgta atatctgtgc tctaataacg 1620
atgtatatat gtatagtgaa catattgctt agtaacaaat gaaggacaag catactacat 1680
aatataatcc atacaattct cttcagttag tgtaaactgc aaatactaca gataattctg 1740
ataataaaat gatatgcacg ctgaatcctg ctatggtcac cattctaatg tatgtagtat 1800
ttcaaatttc cttccttgta actttcaaag aaagccatct tattcttgta aaattttaga 1860
tggtattatc acagatttaa aaaggttgta ttacatattg tttaaacttt gtaagtagaa 1920
atatatctgt tataattata caggctctgt ggagaaataa agttcaaaaa atattaattt 1980
gtaaaatcag ctcgttttaa agtgtgcttg tgttgtcaaa aatatcagat agtaatacac 2040
agtgagcatt tttaaacaaa gggaaaccta tatttatgta actgtatact gaattctgac 2100
aaaataaaaa aagatacctt attgacgaaa tatttaggat aaacaaaatt ctatttaatc 2160
caccttaaaa cctaaatgta ttttcatgga tttcatttgt tggtacatat tacacaaaac 2220
attgtgcctt aaaatgagtc atacatcttt taaattggaa tgcagtaata gatatgtgat 2280
tttacatcat ttttaagaaa ccaaggggaa gtaataagtt gaaaaagaaa tccataacta 2340
ttaaaagatt ttaacttttt tattttatta aaatgcttgc atattttaag taaaaaaaaa 2400
aaaaaaaaaa cc 2412

<210> 85
<211> 1004

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7509439CB1

<400> 85
ggctcgggac gcgcgccatt gtgttggtac ccgggaattc ggccattatg gccgggggac 60
cttcagcagg gctgtggcta ccatgttctc tcgcgcgggt gtcgctgggc tgtcggcctg 120
gaccttgcag ccgcaatgga ttcaagttcg aaatatggca actttgaaag atatcaccag 180
gagactaaag tccatcaaaa acatccagaa aattaccaag tctatgaaaa tggtagcggc 240
agcaaaatat gcccgagctg agagagagct gaaaccagct cgaatatatg gattgggatc 300
tttagctctg tatgaaaaag ctgatatcaa ggggcctgaa gacaagaaga aacacctcct 360
tattggtgtg tcctcagatc gaggactgtg tggtgctatt cattcctcca ttgctaaaca 420
gatgaaaagc gaggttgcta cactaacagc agctgggaaa gaagttatgc ttgttggaat 480
tggtgacaaa atcagaggca tactttatag ttctttgcag gtgctgaagg aaaggaatga 540
tgatagtgtg tggaacaact caggaaatca ccaccaccca taccccaaag acctgatcca 600
tggactcatt ctgaccagtt tctggtggca ttcaaagaag tgggaagaaa gccccccact 660
tttggagatg cgtcagtcat tgcccttgaa ttactaaatt ctggatatga atttgatgaa 720
ggctccatca tctttaataa attcaggtct gtcatctcct ataagacaga agaaaagccc 780
atcttttccc tttaataccg ttgcaagtgc tgacagcatg agtatctatg acgatattga 840
tgctgacgtg ctgcaaatta ccaagaatac aatctggcca acatcatcta ctactctctg 900
aaggagtcca ccactagtga gcagagtgcc aggatgacag ccatggacaa tgccagcagg 960
ctttccacac ctgcgccgta ctatagtccg gtcgtcactt ggtc 1004

<210> 86
<211> 5231
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510202CB1

<400> 86
atgagcaaga gacgcatgag cgtgggtcag caaacatggg ctcttctctg caagaactgt 60
ctcaaaaaat ggagaatgaa aagacagacc ttgttggaat ggctcttttc atttcttctg 120
gtactgtttc tgtacctatt tttctccaat ttacatcaag ttcatgacac tcctcaaatg 180
tcttcaatgg atctgggacg tgtagatagt tttaatgata ctaattatgt tattgcattt 240
gcacctgaat ccaaaactac ccaagagata atgaacaaag tggcttcagc cccattccta 300
aaaggaagaa caatcatggg gtggcctgat gaaaaaagca tggatgaatt ggatttgaac 360
tattcaatag acgcagtgag agtcatcttt actgatacct tctcctacca tttgaagttt 420
tcttggggac atagaatccc catgatgaaa gagcacagag accattcagc tcactgtcaa 480
gcagtgaatg aaaaaatgaa gtgtgaaggt tcagagttct gggagaaagg ctttgtagct 540
tttcaagctg ccattaatgc tgctatcata gaaatcgcaa caaatcattc agtgatggaa 600
cagctgatgt cagttactgg tgtacatatg aagatattac cttttgttgc ccaaggagga 660
gttgcaactg atttttttcat tttcttttgc attatttctt tttctacatt tatatactat 720
gtatcagtca atgttacaca agaaagacaa tacattacgt cattgatgac aatgatggga 780
ctccgagagt cagcattctg gctttcctgg ggtttgatgt atgctggctt catccttatc 840
atggccactt taatggctct tattgtaaaa tctgcacaaa ttgtcgtcct gactggtttt 900
gtgatggtct tcaccctctt tctcctctat ggcctgtctt tgataacttt agctttcctg 960
atgagtgtgt tgataaagaa acctttcctt acgggcttgg ttgtgtttct ccttattgtc 1020
ttttggggga tcctgggatt cccagcattg tatacacatc ttcctgcatt tttggaatgg 1080
actttgtgtc ttcttagccc ctttgccttc actgttggga tggcccagct tatacatttg 1140
gactatgatg tgaattctaa tgcccacttg gattcttcac aaaatccata cctcataata 1200

gctactcttt tcatgttggt ttttgacacc cttctgtatt tggtattgac attatatttt 1260
gacaaaattt tgcccgctga atatggacat cgatgttctc ccttgttttt cctgaaatcc 1320
tgtttttggt ttcaacacgg aagggctaat catgtggtcc ttgagaatga aacagattct 1380
gatcctacac ctaatgactg ttttgaacca gtgtctccag aattctgtgg gaaggaagcc 1440
atcagaatca aaaatcttaa aaaagaatat gcagggaagt gtgagagagt agaagctttg 1500
aaaggtgtgg tgtttgacat atatgaaggc cagatcactg ccctccttgg tcacagtgga 1560
gctggaaaaa ctaccctgtt aaacatactt agtgggttgt cagttccaac atcaggttca 1620
gtcactgtct ataatcacac actttcaaga atggctgata tagaaaatat cagcaagttc 1680
actggatttt gtccacaatc caatgtgcaa tttggatttc tcactgtgaa agaaaacctc 1740
aggctgtttg ctaaaataaa agggattttg ccacatgaag tggagaaaga ggttttgcta 1800
ttggatgaac cgactgctgg attggatcct ctttcaaggc accgaatatg gaatctcctg 1860
aaagagggga aatcagacag agtaattctc ttcagcaccc agtttataga tgaggctgac 1920
attctggcgg acaggaaggt gttcatatcc aatgggaagc tgaagtgtgc aggctcttct 1980
ctgttcctta agaagaaatg ggcataggc taccatttaa gtttgcatct gaatgaaagg 2040
tgtgatccag agagtataac atcactggtt aagcagcaca tctctgatgc caaattgaca 2100
gcacaaagtg aagaaaaact tgtatatatt ttgcctttgg aaaggacaaa caaatttcca 2160
gaactttaca gggatcttga tagatgttct aaccaaggca ttgaggatta tggtgtttcc 2220
ataacaactt tgaatgaggt gtttctgaaa ttagaaggaa aatcaactat tgatgaatca 2280
gatattggaa tttggggaca attacaaact gatggggcaa aagatatagg aagccttgtt 2340
gagctggaac aagttttgtc ttccttccac gaaacaagga aacaatcag tggcgtggcg 2400
ctctggaggc agcaggtctg tgcaatagca aaagttcgct tcctaaagtt aaagaaagaa 2460
agaaaaagcc tgtggactat attattgctt tttggtatta gctttatccc tcaacttttg 2520
gaacatctat tctacgagtc atatcagaaa agttacccgt gggaactgtc tccaaataca 2580
tacttcctct caccaggaca acaaccacag gatcctctga cccatttact ggtcatcaat 2640
aagacagggt caaccattga taacttttta cattcactga ggcgacagaa catagctata 2700
gaagtggatg cctttggaac tagaaatggc acagatgacc catcttacaa tggtgctatc 2760
attgtgtcag gtgatgaaaa ggatcacaga ttttcaatag catgtaatac aaaacggctg 2820
aattgctttc ctgtcctcct ggatgtcatt agcaatggac tacttggaat ttttaattcg 2880
tcagaacaca ttcagactga cagaagcaca ttttttgaag agcatatgga ttatgagtat 2940
gggtaccgaa gtaacacctt cttctggata ccgatggcag cctctttcac tccatacatt 3000
gcaatgagca gcattggtga ctacaaaaaa aaagctcatt cccagctacg gatttcaggc 3060
ctctaccctt ctgcatactg gtttggccaa gcactggtgg atgtttccct gtactttttg 3120
atcctcctgc taatgcaaat aatggattat atttttagcc cagaggagat tatatttata 3180
attcaaaacc tgttaattca aatcctgtgt agtattggct atgtctcatc tcctgttttc 3240
ttgacatatg tgatttcatt catttttcgc aatgggagaa aaaatagtgg catttggtca 3300
tttttcttct taattgtggt catcttctcg atagttgcta ctgatctaaa tgaatatgga 3360
tttctagggc tattttttgg caccatgtta atacctccct tcacattgat tggctctcta 3420
ttcatttttt ctgagatttc tcctgattcc atggattact taggagcttc agaatctgaa 3480
attgtatacc tggcactgct aataccttac cttcattttc tcatttttct tttcattctg 3540
cgatgcctag aaatgaactg caggaagaaa ctaatgagaa aggatcctgt gttcagaatt 3600
tctccaagaa gcaacgctat ttttccaaac ccagaagagc ctgaaggaga ggaggaagat 3660
atccagatgg aaagaatgag aacagtgaat gctatggctg tgcgagactt tgatgagaca 3720
cccgtcatca ttgccagctg tctacggaag gaatatgcag gcaaaaagaa aaattgcttt 3780
tctaaaagga agaaaacaat tgccacaaga aatgtctctt tttgtgttaa aaaaggtgaa 3840
gttataggac tgttaggaca caatggagct ggtaaaagta caactattaa gatgataact 3900
ggagacacaa aaccaactgc aggacaggtg attttgaaag ggagcggtgg aggggaaccc 3960
ctgggcttcc tggggtactg ccctcaggag aatgcgctgt ggcccaacct gacagtgagg 4020
cagcacctgg aggtgtacgc tgccgtgaaa ggtctcagga aaggggacgc aatgatcgcc 4080
atcacacggt tagtggatgc gctcaagctg caggaccagc tgaaggctcc cgtgaagacc 4140
ttgtcagagg gaataaagcg aaaggtacgg gcagggcttg ttgttgctct gcaagtgccg 4200
tagatgggaa ccagaggggc tgttcctccc gttgcagctg tgctttgtgc tgagcatcct 4260
ggggaacccg tcagtggtgc ttctggatga gccgtcgacc gggatggacc ccgaggggca 4320
gcagcaaatg tggcaggtga ttcgggccac ctttagaaac acggagaggg gcgccctcct 4380
gaccacccac tacatggcag aggctgaggc ggtgtgtgac cgagtggcca tcatggtgtc 4440
aggaaggctg agatgtattg gttccatcca acacctgaaa agcaaatttg gcaaagacta 4500
cctgctggag atgaagctga agaacctggc acaaatggag cccctccatg cagagatcct 4560

gaggcttttc ccccaggctg ctcagcagga aaggttctcc tccctgatgg tctataagtt 4620
gcctgttgag gatgtgcgac ctttatcaca ggctttcttc aaattagaga tagttaaaca 4680
gagtttcgac ctggaggagt acagcctctc acagtctacc ctggagcagg ttttcctgga 4740
gctctccaag gagcaggagc tgggtgatct tgaagaggac tttgatccct cggtgaagtg 4800
gaaactcctc ctgcaggaag agccttaaag ctccaaatac cctatatctt tctttaatcc 4860
tgtgactctt ttaaagataa tattttatag ccttaatatg ccttatatca gaggtggtac 4920
aaaatgcatt tgaaactcat gcaataatta tcctcagtag tatttcttac agtgagacaa 4980
caggcgatgt cagtgagggc gatcataggg cataagccta agccatacca tgcagccttt 5040
gtgccagcaa ccaaatccca tgtttcctac tgtgttaagt ttaaaaatgc atttattata 5100
gaattgtcta catttctgag gatgtcatgg agaatgctta attttcttcc tctgaacttc 5160
aaaatattaa atattttctt atttattttt ttgattaaag tataaattaa gacaccctat 5220
tgacttccgg g 5231

<210> 87
<211> 3269
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510203CB1

<400> 87
ggctttgtga ctgggtgctc ttgggaactt aaaggcccag gtctctgaga acattctcat 60
atcttcagag aagagaaagg agtcaggttt agaatggtct ttttgtaaaa acaaaaattt 120
ttaaactctg aaagattttt atgcagaggg atgtgaccac ctgcaagttt actgggagag 180
aaataggagg aaatgaagct ccagccagat attaggaatt gagcgaagca tgagaaaaag 240
ctctgcagta atgagaagta accctggaca gagtggcagg catagctctg gcggtcctct 300
tggggatgga cattcgtccc tctctgctgg tcggaatcaa ggtttacaag gatggacttg 360
agtcctgctc gaggtctagg ggtatccaga gcagggtggg ttagagaggg aggcctaaga 420
gtcctgctcg aggtctaggg gtatcctgct agggtgggtt agacggggag gcctgagagg 480
atgagaggtg ggatctgcac acgcactgag aaaggcatag tatagactca gaacagtcct 540
ctcccaatct ccccttctac cctccaggac ctccctctga gattctgcca ctggtacaag 600
ctgtccccaa agcctgggct gagaatggac aagagtctga ctcagccaca gcccagtcag 660
tacagaaccc agagaaaaca aaggaggggc tggaggagga gcagagcaca tctggtcgcc 720
tgctgcagga agaaagcaag aaggagggcg ccgtggcctt gcacgtgtac caagcttact 780
ggaaggccgt gggccagggc ttggccttag ccatcctctt ctctctgctt ctcatgcaag 840
ccacgcggaa cgctgctgac tggtggctct cccactggat ctctcagctg aaggctgaga 900
atagctccca ggaggcgcaa ccctccacca gcccagcttc tatggggctc ttctctccgc 960
agctgctcct cttttcccct ggaaacctct acatcccagt gttcccactg cccaaagctg 1020
cccccaatgg ctcctcagac atccgtttct acctcaccgt gtatgcgacc attgctggtg 1080
taaattccct ctgcaccctt ctccgggcag tgctctttgc agcaggcacc cttcaagcag 1140
ctgccactct gcatcgccgc ctgctgcatc gagtccttat ggcaccagtg actttcttca 1200
atgccacacc cacgggccgg atcctaaacc gcttctcctc tgatgtggcc tgtgcggatg 1260
acagcctgcc cttcatcctc aacatcctcc tggccaacgc ggcaggcctg ctggggctcc 1320
tggccgtgct gggctctggc ctgccctggc tgctgctcct gctgccgcct ttgagcatca 1380
tgtactatca cgtgcagcgc cactacaggg cctcctcacg ggagctgcgg cgcctgggca 1440
gcctcaccct gtctccactg tatagccatc tggccgatac cttggctggc ctctctgtgc 1500
tccgggccac aggggccacc tacaggtttg aggaggagaa cctgcgactc cttgagctaa 1560
accagaggtg ccagtttgcc accagtgcca caatgcagtg gctggacatt cggctacagc 1620
tcatgggggc ggcagtggtc agcgctatcg caggcatcgc tctggtgcag caccagcagg 1680
gcctcgctaa cccagggctg gtgggcttgt cgctgtctta tgccctgtcc ctgacgggcc 1740
tgctctcggg cctggtgagc agcttcacac agacagaggc catgctggtg agcgtcgagc 1800
ggctggaaga gtacacctgt gacctgcccc aggaacccca gggccagcca ctgcaggtgg 1860
gcctgtaccc ccaccccagg ccaaagctct ggaaccctga aggccccagt ctccctcaca 1920
attcctttct ttttgcccac ccatctttct cagctcccat aacctctctt catgatgacc 1980

acaattcttc accatgtccc ttcttcccca tctctcattc tctcattcct ctcacactgt 2040
ccatttctca ttattctccc ctcctcacca tcgctcctca tctcccctat ctccctttcc 2100
ccgtctgtct cccacccatg gaccccacca gctgggcacc ggctggctga cccagggggg 2160
cgtggagttc caggacgtgg tgttggcgta ccggccaggg ctgccgaatg ccctggatgg 2220
agtgaccttc tgcgtgcagc ctggagagaa gttgggcatc gtgggccgca caggctccgg 2280
caagtcttcc ctgttgttgg tgctcttccg gctgctagag cccagttcag ggcgagtgct 2340
gctggacggc gtggacacca gccagctgga gctggcccag ctcagatccc agttggctat 2400
catcccccag gagccctttt tgttcagtgg gactgttcgg gaaaacctgg acccccaggg 2460
cctacataag gacagggcct tgtggcaggc cctgaagcag tgccacctga gtgaggtgat 2520
tacatccatg ggtggtctgg atggtgagct gggtgagggg ggccggagct tatctcttgg 2580
gcagaggcag ctgttgtgtt tggccagggc tctcctcaca gatgccaaga tcctgtgtat 2640
cgatgaggcc acagcaagtg tggaccagaa gacagaccag ctgctccagc agaccatctg 2700
caaacgcttt gccaacaaga cagtgctgac cattgcccat aggctcaaca cgatcctgaa 2760
ctcagaccgg gtgctggtgc tacaagcggg gagagtggta gagctggact ccccggccac 2820
cctgcgcaac cagcccccact ccctgttcca gcagctgctg cagagcagcc agcagggagt 2880
ccctgcctca ctcggaggtc cctgagccca atcccacacc ctgcagagtt ctcccctctc 2940
tctgatccag gccgggccta tacagaggtg ctggctgctt gtttacattc tcctctgggg 3000
ctctacctct ccacacttcc ccagaaggga aaagggcacc ctggattact ctttggaaat 3060
cactccttgg tgggcagcat cctgaggctt ccccagaacc aggcctctgc tctggccctc 3120
ttgcatctgg aacgccaggt gggtttttct ggcataggag cccacttgca ttttcatagt 3180
tttatttgat aaaattccat cttacattct gtgtattaaa aaaataatat ttctggtgtg 3240
agaaaaaaaa aaaaaaaaaa atggttcgg 3269

<210> 88
<211> 7706
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510208CB1

<400> 88
gcgccccgcc cccgcgcggg cgatgcccag cggcgcggcg ggctgcgggg cccggcgggg 60
cgcgcagagg agcgggccgc ggcgctgagg cggcggagcg tggccccgcc atgggcttcc 120
tgcaccagct gcagctgctg ctctggaaga acgtgacgct caaacgccgg agcccgtggg 180
tcctggcctt cgagatcttc atccccctgg tgctgttctt tatcctgctg ggctgcgac 240
agaagaagcc caccatctcc gtgaaggaag tctccttcta cacagcggcg cccctgacgt 300
ctgccggcat cctgcctgtc atgcaatcgc tgtgcccgga cggccagcga gacgagttcg 360
gcttcctgca gtacgccaac tccacggtca cgcagctgct tgagcgcctg gaccgcgtgg 420
tggaggaagg caacctgttt gacccagcgc ggcccagcct gggctcagag ctcgaggccc 480
tacgccagca tctggaggcc ctcagtgcgg gcccgggcac ctcggggagc cacctggaca 540
gatccacagt gtcttccttc tctctggact cggtggccag aaacccgcag gagctctggc 600
gtttcctgac gcaaaacttg tcgctgccca atagcacggc ccaagcactc ttggccgccc 660
gtgtggaccc gcccgaggtc taccacctgc tctttggtcc ctcatctgcc ctggattcac 720
agtctggcct ccacaagggt caggagccct ggagccgcct agggggcaat cccctgttcc 780
ggatggagga gctgctgctg gctcctgccc tcctggagca gctcacctgc acgccgggct 840
cgggggagct gggccggatc ctcactgtgc ctgagagtca gaagggagcc ctgcagggct 900
accgggatgc tgtctgcagt gggcaggctg ctgcgcgtgc caggcgcttc tctgggctgt 960
ctgctgagct ccggaaccag ctggacgtgg ccaaggtctc ccagcagctg ggcctggatg 1020
cccccaacgg ctcggactcc tcgccacagg cgccaccccc acggaggctg caggcgcttc 1080
tgggggacct gctggatgcc cagaaggttc tgcaggatgt ggatgtcctg tcggccctgg 1140
ccctgctact gccccagggt gcctgcactg gccggacccc cggaccccca gccagtggtg 1200
cgggtggggc ggccaatggc actggggcag gggcagtcat gggccccaac gccaccgctg 1260
aggagggcgc accctctgct gcagcactgg ccaccccgga cacgctgcag ggccagtgct 1320
cagccttcgt acagctctgg gccggcctgc agcccatctt gtgtggcaac aaccgcacca 1380

ttgaacccga ggcgctgcgg cggggcaaca tgagctccct gggcttcacg agcaaggagc 1440
agcggaacct gggcctcctc gtgcacctca tgaccagcaa ccccaaaatc ctgtacgcgc 1500
ctgcgggctc tgaggtcgac cgcgtcatcc tcaaggccaa cgagactttt gcttttgtgg 1560
gcaacgtgac tcactatgcc caggtctggc tcaacatctc ggcggagatc cgcagcttcc 1620
tggagcaggg caggctgcag caacacctgc gctggctgca gcagtatgta gcagagctgc 1680
ggctgcaccc cgaggcactg aacctgtcac tggatgagct gccgccggcc ctgagacagg 1740
acaacttctc gctgcccagt ggcatggccc tcctgcagca gctggatacc attgacaacg 1800
cggcctgcgg ctggatccag ttcatgtcca aggtgagcgt ggacatcttc aagggcttcc 1860
ccgacgagga gagcattgtc aactacaccc tcaaccaggc ctaccaggac aacgtcactg 1920
tttttgccag tgtgatcttc cagacccgga aggacggctc gctcccgcct cacgtgcact 1980
acaagatccg ccagaactcc agcttcaccg agaaaaccaa cgagatccgc cgcgcctact 2040
ggcggcctgg gcccaatact ggcggccgct tctacttcct ctacggcttc gtctggatcc 2100
aggacatgat ggagcgcgcc atcatcgaca cttttgtggg gcacgacgtg gtggagccag 2160
gcagctacgt gcagatgttc ccctacccct gctacacacg cgatgacttc ctgtttgtca 2220
ttgagcacat gatgccgctg tgcatggtga tctcctgggt ctactccgtg gccatgacca 2280
tccagcacat cgtggcggag aaggagcacc ggctcaagga ggtgcgcggg ccgggactct 2340
ccctggaggc gcgggctggc agggaggggc ggcgtcctcc ccggggcctg ccccaagccc 2400
ctggcccacc cgcaggtgat gaagaccatg ggcctgaaca acgcggtgca ctgggtggcc 2460
tggttcatca ccggctttgt gcagctgtcc atctccgtga cagcactcac cgccatcctg 2520
aagtacggcc aggtgcttat gcacagccac gtggtcatca tctggctctt cctggcagtc 2580
tacgcggtgg ccaccatcat gttctgcttc ctggtgtctg tgctgtactc caaggccaag 2640
ctggcctcgg cctgcggtgg catcatctac ttcctgagct acgtgcccta catgtacgtg 2700
gcgatccgag aggaggtggc gcatgataag atcacggcct tcgagaagtg catcgcgtcc 2760
ctcatgtcca cgacggcctt tggtctgggc tctaagtact tcgcgctgta tgaggtggcc 2820
ggcgtgggca tccagtggca caccttcagc cagtccccgg tggaggggga cgacttcaac 2880
ttgctcctgg ctgtcaccat gctgatggtg gacgccgtgg tctatggcat cctcacgtgg 2940
tacattgagg ctgtgcaccc aggcatgtac gggctgcccc ggccctggta cttcccactg 3000
cagaagtcct actggctggg cagtgggcgg acagaagcct gggagtggag ctggccgtgg 3060
gcacgcaccc cccgcctcag tgtcatggag gaggaccagg cctgtgccat ggagagccgg 3120
cgctttgagg agacccgtgg catggaggag gagcccaccc acctgcctct ggttgtctgc 3180
gtggacaaac tcaccaaggt ctacaaggac gacaagaagc tggccctgaa caagctgagc 3240
ctgaacctct acgagaacca ggtggtctcc ttcttgggcc acaacggggc gggcaagacc 3300
accaccatgt ccatcctgac cggcctgttc cctccaacgt cgggttccgc caccatctac 3360
gggcacgaca tccgcacgga gatggatgag atccgcaaga acctgggcat gtgcccgcag 3420
cacaatgtgc tctttgaccg gctcacggtg gaggaacacc tctggttcta ctcacggctc 3480
aagagcatgg ctcaggagga gatccgcaga gagatggaca agatgatcga ggacctggag 3540
ctctccaaca aacggcactc actggtgcag acattgtcgg gtggcatgaa gcgcaagctg 3600
tccgtggcca tcgccttcgt gggcggctct cgcgccatca tcctggacga gcccacggcg 3660
ggcgtggacc cctacgcgcg ccgcgccatc tgggacctca tcctgaagta caagccaggc 3720
cgcaccatcc ttctgtccac ccaccacatg gatgaggctg acctgcttgg ggaccgcatt 3780
gccatcatcg tcccatggga agctcaagtg ctgcggctcc ccgctcttcc tcaagggcac 3840
ctatggcgac gggtaccgcc tcacgctggt caagcggccc gccgagccgg ggggccccca 3900
agagccaggg ctggcatcca gcccccccagg tcgggccccg ctgagcagct gctccgagct 3960
ccaggtgtcc cagttcatcc gcaagcatgt ggcctcctgc ctgctggtct cagacacaag 4020
cacggagctc tcctacatcc tgcccagcga ggccgccaag aaggggggctt tcgagcgcct 4080
cttccagcac ctggagcgca gcctggatgc actgcacctc agcagcttcg ggctgatgga 4140
cacgaccctg gaggaagtgt tcctcaaggt gtcggaggag gatcagtcgc tggagaacag 4200
tgaggccgat gtgaaggagt ccaggaagga tgtgctccct ggggcggagg gcccggcgtc 4260
tggggagggt cacgctggca atctggcccg gtgctcggag ctgacccagt cgcaggcatc 4320
gctgcagtcg gcgtcatctg tgggctctgc ccgtggcgac gagggagctg gctacaccga 4380
cgtctatggc gactaccgcc ccctctttga taacccacag gacccagaca atgtcagcct 4440
gcaagaggtg gaggcagagg ccctgtcgag ggtcggccag ggcagccgca agctggacgg 4500
cgggtggctg aaggtgcgcc agttccacgg gctgctggtc aaacgcttcc actgcgcccg 4560
ccgcaactcc aaggcactct tctcccagat cttgctgcca gccttcttcg tctgcgtggc 4620
catgaccgtg gccctgtccg tcccggagat tggtgatctg cccccgctgg tcctgtcacc 4680
ttcccagtac cacaactaca cccagccccg tggcaatttc atcccctacg ccaacgagga 4740

gcgccgcgag taccggctgc ggctatcgcc cgacgccagc ccccagcagc tcgtgagcac 4800
gttccggctg ccgtcggggg tgggtgccac ctgcgtgctc aagtctcccg ccaacggctc 4860
gctggggccc acgttgaacc tgagcagcgg ggagtcgcgc ctgctggcgg ctcggttctt 4920
cgacagcatg tgtctggagt ccttcacaca ggggctgcca ctgtccaatt tcgtgccacc 4980
cccaccctcg cccgccccat ctgactcgcc agcgtccccg gatgaggacc tgcaggcctg 5040
gaacgtctcc ctgccgccca ccgctgggcc agaaatgtgg acgtcggcac cctccctgcc 5100
gcgcctggta cgggagcccg tccgctgcac ctgctctgcg cagggcaccg gcttctcctg 5160
ccccagcagt gtgggcgggc acccgcccca gatgcgggtg gtcacaggcg acatcctgac 5220
cgacatcacc ggccacaatg tctctgagta cctgctcttc acctccgacc gcttccgact 5280
gcaccggtat ggggccatca cctttggaaa cgtcctgaag tccatcccag cctcatttgg 5340
caccagggcc ccacccatgg tgcggaagat cgcggtgcgc agggctgccc aggttttcta 5400
caacaacaag ggctatcaca gcatgcccac ctacctcaac agcctcaaca acgccatcct 5460
gcgtgccaac ctgcccaaga gcaagggcaa cccggcggct tacggcatca ccgtcaccaa 5520
ccaccccatg aataagacca gcgccagcct ctccctggat tacctgctgc agggcacgga 5580
tgtcgtcatc gccatcttca tcatcgtggc catgtccttc gtgccggcca gcttcgttgt 5640
cttcctcgtg gccgagaagt ccaccaaggc caagcatctg cagtttgtca gcggctgcaa 5700
ccccatcatc tactggctgg cgaactacgt gtgggacatg ctcaactacc tggtccccgc 5760
tacctgctgt gtcatcatcc tgtttgtgtt cgacctgccg gcctacacgt cgcccaccaa 5820
cttccctgcc gtcctctccc tcttcctgct ctatgggtgg tccatcacgc ccatcatgta 5880
cccggcctcc ttctggttcg aggtccccag ctccgcctac gtgttcctca ttgtcatcaa 5940
tctcttcatc ggcatcaccg ccaccgtggc caccttcctg ctacagctct tcgagcacga 6000
caaggacctg aaggttgtca acagttacct gaaaagctgc ttcctcattt tccccaacta 6060
caacctgggc cacgggctca tggagatggc ctacaacgag tacatcaacg agtactacgc 6120
caagattggc cagtttgaca agatgaagtc cccgttcgag tgggacattg tcacccgcgg 6180
actggtggcc atggcggttg agggcgtcgt gggcttcctc ctgaccatca tgtgccagta 6240
caacttcctg cggcggccac agcgcatgcc tgtgtctacc aagcctgtgg aggatgatgt 6300
ggacgtggcc agtgagcggc agcgagtgct ccgggggagac gccgacaatg acatggtcaa 6360
gattgagaac ctgaccaagg tctacaagtc ccggaagatt ggccgtatcc tggccgttga 6420
ccgcctgtgc ctgggtgtgc gtcctggcga gtgcttcggg ctcctgggcg tcaacggtgc 6480
gggcaagacc agcaccttca agatgctgac cggcgacgag agcacgacgg ggggcgaggc 6540
cttcgtcaat ggacacagcg tgctgaagga gctgctccag gtgcagcaga gcctcggcta 6600
ctgcccgcag tgtgacgcgc tgttcgacga gctcacggcc cgggagcacc tgcagctgta 6660
cacgcggctg cgtgggatct cctggaagga cgaggcccgg gtggtgaagt gggctctgga 6720
gaagctggag ctgaccaagt acgcagacaa gccggctggc acctacagcg gcggcaacaa 6780
gcggaagctc tccacggcca tcgccctcat tgggtaccca gccttcatct tcctggacga 6840
gcccaccaca ggcatggacc ccaaggcccg gcgcttcctc tggaacctca tcctcgacct 6900
catcaagaca gggcgttcag tggtgctgac atcacacagc atggaggagt gcgaggcgct 6960
gtgcacgcgg ctggccatca tggtgaacgg tcgcctgcgg tgcctgggca gcatccagca 7020
cctgaagaac cggtttggag atggctacat gatcacggtg cggaccaaga gcagccagag 7080
tgtgaaggac gtggtgcggt tcttcaaccg caacttcccg gaagccatgc tcaaggagcg 7140
gcaccacaca aaggtgcagt accagctcaa gtcggagcac atctcgctgg cccaggtgtt 7200
cagcaagatg gagcaggtgt ctggcgtgct gggcatcgag gactactcgg tcagccagac 7260
cacactggac aatgtgttcg tgaactttgc caagaagcag agtgacaacc tggagcagca 7320
ggagacggag ccgccatccg cactgcagtc cccctctcgg ctgcttgctc agcctgctcc 7380
ggccccggtc tgcccccacg gagctccggg cacttgtggc agacgagccc gaggacctgg 7440
acacggagga cgagggcctc atcagcttcg aggaggagcg ggcccagctg tccttcaaca 7500
cggacacgct ctgctgacca cccagagctg ggccagggag gacacgctcc actgaccacc 7560
cagagctggg ccagggactc aacaatgggg acagaagtcc cccagtgcct gccagggcct 7620
ggagtggagg ttcaggacca aggggcttct ggtcctccag cccctgtact cggccatgcc 7680
ctgcggtcac tgcggttgcc ggccct 7706

<210> 89
<211> 3159
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510446CB1

<400> 89
agcaagagca gaggcttaag gagctacact gggggaagga caggggcaag caggccaagg 60
cctggccggg gctcgggggg agggaatatg gagcaatccc ggtcacagca gcgtgggggt 120
gaacaaagct ggtggggtag tgacccccag taccagtata tgccctttga acactgcacc 180
agctacggac tgccctctga gaatggggc ctccagcaca ggctccggaa ggatgcaggc 240
ccccgccaca acgtccaccc cacacagatt tatggccatc acaaagaaca attctcagac 300
agggagcagg acatagggat gcccaagaag acaggctcca gttctaccgt ggacagcaag 360
gatgaggatc actattctaa atgtcaaggt gatggggact gaggaataaa gaaatctgga 420
gtagaaacag attgtatcca ccgcctggga caggtggtga gaagaaaatt aggggaagac 480
tggatctttc tggtgcttct gggactgctg atggctctgg tcagctggag catggactac 540
gtcagtgcca aaagccttca ggcctacaag tggtcctacg cgcagatgca gcccagcctt 600
cctctgcagt tcctggtctg ggtcaccttc ccactagtcc tcatcctctt cagcgccctc 660
ttctgccacc tcatctctcc ccaggctgtt ggctctggaa tccccgaaat gaagacaata 720
cttcgtgggg ttgtcctgaa ggaatacctc acaatgaaag cctttgtggc caaggttgtc 780
gccctgactg cgggcctggg cagtggcatc cccgtgggga aagagggccc cttcgtccac 840
attgccagca tctgtgctgc tgtcctcagc aaattcatgt ctgtgttctg cggggtatat 900
gagcagccat actactactc tgatatcctg acggtgggct gtgctgtggg agtcggctgt 960
tgttttggga caccacttgg aggagtgcta tttagcatcg aggtcacctc cacctacttt 1020
gctgttcgga actactggag aggattcttt gcagccacgt tcagcgcctt tgtgtttcga 1080
gtgctggcag tgtggaacaa ggatgctgtc accatcactg ctctgttcag aaccaatttc 1140
cgaatggatt tcccctttga cctgaaggaa ctaccagctt ttgctgccat cgggatttgc 1200
tgtgggctcc tgggagctgt atttgtgtat ctgcatcgcc aagtcatgct cggtgtccga 1260
aagcacaagg ccctcagcca gtttcttgct aagcaccgcc tgctgtatcc tggaattgtt 1320
acctttgtca ttgcctcatt caccttccca ccaggaatgg gtcaattcat ggctggagag 1380
ttgatgcccc gcgaagccat cagtactttg tttgacaaca atacatgggt gaaacacgcg 1440
ggtgatcctg agagcctggg ccagtcagct gtgtggattc acccccgggt caacgttgtc 1500
atcatcatct ttctcttctt cgtcatgaag ttctggatgt ccatcgtggc caccactatg 1560
cccatacccct gcggaggctt catgcctgtg tttgtgctag gagctgcatt tggaaggctg 1620
gtaggagaaa tcatggccat gctctttcct gatggtattt tgtttgatga catcatctac 1680
aagatcctac ctgggggcta tgcagtaatt ggagcagcag cgctgactgg tgccgtttcc 1740
cacacagtct ccacagctgt gatttgcttc gaattaacgg gtcagattgc tcacatcctg 1800
cccatgatgg tggctgttat cttggccaac atggtggccc agagcctgca gccctctctc 1860
tatgacagca tcatccaggt caagaagcta ccctacttgc ctgaccttgg ctggaaccag 1920
ctcagcaaat ataccatctt tgttgaggac atcatggtac gtgatgtgaa gtttgtttca 1980
gcttcttaca catatgggga gttgcgaacc ctgctccaga ccaccacagt caagacttta 2040
ccactggttg actcaaaaga ttcaatgatc ctgctgggct cggtggagcg gtcggaactg 2100
caggccctcc tgcagcgcca cctgtgtcct gagcgcagac tgcgcgcagc ccaagagatg 2160
gcgcggaagt tgtcggagct gccttacgac gggaaggcgc ggctggctgg ggaggggctc 2220
cccggcgcgc ctccaggccg gcccgagtcc ttcgcctttg tggatgagga tgaggacgaa 2280
gacctctctg gcaagagcga gcttcctctt tcccttgctc tccacccctc tactactgcc 2340
cctctgtccc cagaagagcc caatgggcct ctgcctggcc acaaacagca gccggaagca 2400
ccagagcctg caggtcaaag accctccatc ttccagtccc tgcttcactg cttgctgggc 2460
agagctcgcc ccacaaagaa gaaaacaacc caggattcca cagatttagt ggataacatg 2520
tcacctgaag agattgaggc ctgggagcag gagcagctga gccagcctgt ctgttttgat 2580
tcctgctgta ttgaccagtc tcccttccag ctggtggagc agacaacccct gcacaagact 2640
catacccctgt tttcactcct tggcctccac ctcgcttacg tgaccagcat ggggaagctc 2700
aggggcgtcc tggccctgga ggagctacag aaggccattg aggggcacac caagtctggg 2760
gtgcagctcc gccctcccct tgccagcttc cggaacacga cttcaactcg aaagagtacc 2820
ggggcacctc catcttctgc agagaactgg aacctgcctg aggacaggcc tggggccact 2880
ggaacagggg atgtgattgc tgcctcccca gagacccctg tgccatctcc ttccccagag 2940
ccccctctct ccctggcccc aggcaaggta gagggcgagt tggaggagct ggagctggtg 3000
gagagtccag ggctggaaga ggagctggcc gacatcttgc agggccccag cctgcgatcc 3060

acagacgagg aggatgagga tgaactgatc ctttgacccc ctcccacgac ctcctcataa 3120
agaccgtgga gaggcccaaa aaaaaaaaaa aaaaaaaaa 3159

<210> 90
<211> 1821
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505294CB1

<400> 90
gctgcctgcc ggtgctcttc gtggctctgg gcatggcctc ggaccccatc ttcacgctgg 60
cgcccccgct gcattgccac tacggggcct tccccccctaa tgcctctggc tgggagcagc 120
ctcccaatgc cagcggcgtc agcgtcgcca gcgctgccct agcagccagc gccgccagcc 180
gtgtcgccac cagtaccgac ccctcgtgca gcggcttcgc cccgccggac ttcaaccatt 240
gcctcaagga ttgggactat aatggccttc ctgtgctcac caccaacgcc atcggccagt 300
gggatctggt gtgtgacctg ggctggcagg tgatcctgga gcagatcctc ttcatcttgg 360
gctttgcctc cggctacctg ttcctgggtt accccgcaga cagatttggc cgtcgcggga 420
ttgtgctgct gaccttgggg ctggtgggcc cctgtggagt aggaggggct gctgcaggct 480
cctccacagg cgtcatggcc ctccgattcc tcttgggctt tctgcttgcc ggtgttgacc 540
tgggtgtcta cctgatgcgc ctggagctgt gcgacccaac ccagaggctt cgggtggccc 600
tggcagggga gttggtgggg gtgggagggc acttcctgtt cctgggcctg gcccttgtct 660
ctaaggattg gcgattccta cagcgaatga tcaccgctcc ctgcatcctc ttcctgtttt 720
atggctggcc tggtttgttc ctggagtccg cacggtggct gatagtgaag cggcagattg 780
aggaggctca gtctgtgctg aggatcctgg ctgagcgaaa ccggccccat gggcagatgc 840
tgggggagga ggcccaggag gccctgcagg cttcattgcc catgccattc gccactgcta 900
ccagcctgtg ggaggaggag ggagcccatc ggacttctac ctgtgctctc tgctggccag 960
cggcaccgca gccctggcct gtgtcttcct gggggtcacc gtggaccgat ttggccgccg 1020
gggcatcctt cttctctcca tgacccttac cggcattgct tccctggtcc tgctgggcct 1080
gtgggattat ctgaacgagg ctgccatcac cactttctct gtccttgggc tcttctcctc 1140
ccaagctgcc gccatcctca gcaccctcct tgctgctgag gtcatcccca ccactgtccg 1200
gggccgtggc ctgggcctga tcatggctct aggggcgctt ggaggactga gcggcccggc 1260
ccagcgcctc cacatgggcc atggagcctt cctgcagcac gtggtgctgg cggcctgcgc 1320
cctcctctgc attctcagca ttatgctgct gccggagacc aagcgcaagc tcctgcccga 1380
ggtgctccgg gacggggagc tgtgtcgccg gccttccctg ctgcggcagc cacccccctac 1440
ccgctgtgac cacgtcccgc tgcttgccac ccccaaccct gccctctgag cggcctctga 1500
gtaccctggc gggaggctgg cccacacaga aaggtggcaa gaagatcggg aagactgagt 1560
agggaaggca gggctgccca gaagtctcag aggcacctca cgccagccat cgcggagagc 1620
tcagagggcc gtccccaccc tgcctcctcc ctgctgcttt gcattcactt ccttggccag 1680
agtcagggga cagggagaga gctccacact gtaaccactg ggtctgggct ccatcctgcg 1740
cccaaagaca tccacccaga cctcattatt tcttgctcta tcattctgtt tcaataaaga 1800
catttggaat aaacgagcaa a 1821

<210> 91
<211> 3526
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505631CB1

<400> 91
cacttggcgg ccatggcagc tgtagtatcg gcgactccgg gtcaaggccc ggtcgagtgc 60

agtaccatgg gcagcaccgg gtatagggca gagacagctt tgtgtcaact ttgctgctga 120
acccctagga cccatcgtta gagacctgca ggactccttt cctcatccca ggctcggagg 180
agagtttgct gggactggtg ggctggtttc ctgctctggg gggcggatca ccttcggggc 240
cgcctcttgg agacaggggc gcctagggaa cgaacaggtt cgcttgagtc acttacccgc 300
cgccgcctaa gacattgtgc caccctcaat ccgacaatcg aagaaatcga tcattcgcac 360
attttcccca ttgacttttc ccatctctgt taacccacga gaatctaatg actggcatct 420
gagaacccag agcctgggac cttagattgc tgtaagcttt ctctggtgct aatatcagca 480
aaaagggtct gttgccgggt acgttcaaga ggaaggtgcc tcgtgaacac atctgctggt 540
gggaaggcct aaagaactgg aaagcccact ctcttggaac caccacacct gtttaaagaa 600
cctaagcacc atttaaagcc actggaaatt tgttgtctag tggttgtggg tgaataaagg 660
agggcagaat ggatgatttc atctccatta gcctgctgtc tctggctatg ttggtgggat 720
gttacgtggc cggaatcatt cccttggctg ttaatttctc agaggaacga ctgaagctgg 780
tgactgtttt gggtgctggc cttctctgtg gaactgctct ggcagtcatc gtgcctgaag 840
gagtacatgc cctttatgaa gatattcttg agggaaaaca ccaccaagca agtgaaacac 900
ataatgtgat tgcatcagac aaagcagcag aaaaatcagt tgtccatgaa catgagcaca 960
gccacgacca cacacagctg catgcctata ttggtgtttc cctcgttctg ggcttcgttt 1020
tcatgttgct ggtggaccag attggtaact cccatgtgca ttctactgac gatccagaag 1080
cagcaaggtc tagcaattcc aaaatcacca ccacgctggg tctggttgtc catgctgcag 1140
ctgatggtgt tgctttggga gcagcagcat ctacttcaca gaccagtgtc cagttaattg 1200
tgtttgtggc aatcatgcta cataaggcac cagctgcttt tggactggtt tccttcttga 1260
tgcatgctgg cttagagcgg aatcgaatca gaaagcactt gctggtcttt gcattggcag 1320
caccagttat gtccatggtg acatacttag gactgagtaa gagcagtaaa gaagcccttt 1380
cagaggtgaa cgccacggga gtggccatgc ttttctctgc cgggacattt ctttatgttg 1440
ccacagtgag gaaagtagca caaataggat acagttgtat gtagtcattg gcaacaattg 1500
catacaattt tactaccaag agaaggtata gtatggaaag tccaaatgac ttccttgatt 1560
ggatgttaac agctgactgg tgtgagactt gaggtttcat ctagtccttc aaaactatat 1620
ggttgcctag attctctctg gaaactgact ttgtcaaata aatagcagat tgtagtgtct 1680
ggtttggttt ggacagtagt gctttctatc atattgttgt gtgcaatggt aatttgttct 1740
actggccaaa gcctctttca gcagtgcctt gccatcatgc ttaaaagttt ggctagtata 1800
tcttgctgga tggagccttg aactccggca aggattgaac catctgactt ccaaatttgc 1860
cttcccctct ggacctcact attaacaagc aaacctttca gggccctctt agctctcaga 1920
agctatgtat gggctttccc agattttaaa gctgctgcct cgagaactac tcatttctct 1980
cctggtcagc agacagaaat agccatacta atctcatagg gctcaaatgc atcttcaggc 2040
agcagggaac caagcagcgt ggcacaggcc ttcttgactg gaggaagagc ttgctggcat 2100
ggtgggcagt attccaggag aggccatgtc cgtgttcact tcttggcaca tttcagttcc 2160
gttttcctct tgtttaaaac tgcctcttta gatgtggatg ccttaatgct gtaacacatt 2220
tgaaaacatt ggcaatactt aagttgctgc catgattaca gatggaatta ttggctacca 2280
aagagacgca attgatgatg agaagcatga ttcttgcttc catataacca aagttaatct 2340
taattgcaat ttgactccgt ttccttggta gggatagact ttcttcagat tccaagtgct 2400
ctcttaaatg gcaaattaag ttaaagaata ctactgctcc attcccctca cttattctcc 2460
agttaattgc ttgtcagttc catttcaaga aagcagtgat gttccaggtt tgattcagtt 2520
ttcctgtgca cactattgcc aaattttttt ttagcaaaga ttctgcactg gaacgtagac 2580
agttggaaac agtactacct acctagaggt tatgtgtttt ctctttctcc ccgctttcac 2640
ctctttcttt cccaattcaa aacagccaag tgagccctgt tctggtattt tgaatcatta 2700
gagaaaagaa agggagtggc tgttttgagt tgtcctttct ttgcagaaag gagaaaatgt 2760
gattgtgttt ttttttaccc agcctacttc taagtgtcac tgcctggttt ttctcttttt 2820
caaggattag aactaagagg acacaccagc atcggagtgt attaagcccc tgaaacacat 2880
ggtagctagg gactgaacac aggaaccgta tgacagcagc acaaaccccc aaaggatgtt 2940
cctgccttgt gggcccctga gccccttggg agactgagaa tcatgaccag attcatccag 3000
aactgctgca gtgttaagtg aaaatcctct gtagttgttc tgcagaggaa ccttccttcc 3060
attagaaaat ttctgctcaa tacagaatgg tccacatcac ccaaagtgca ctgttggaga 3120
tgctgtgaaa ttaaaacctc tttgtacctg agacatctag attcacctca ggaggcctga 3180
aggaaatgtg taacttgtgg gaaagaacta gacaaccatt taggaattct ctagatatac 3240
tcagcctaac ccagtggctt aacacaagga gattggcttt gatctttttt tcttgtggca 3300
tcttccagca agttagaagt ctcatgggat aagactgcag ttcccctggt tcaatagctg 3360
gaacagtgat tttaaatgtc ccttttctg gatcccttgt aaacatgaaa tcattccatg 3420

gatggctgcc ttataatttt gtctctttcc actttaattg tgaatggtta aaaaaatgct 3480
gttttctgat attaaatttt tattagtgca taccttaaaa aaaaaa 3526

<210> 92
<211> 946
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506561CB1

<400> 92
cacggagatt cggcacgagg aggaaaagca gcaaaccaga tggagaagga aggccaaaaa 60
gataagtaga aagctggacc atgagctcct tgagggcagg gactgaatgg ttactatgtc 120
cccagggccc agcatgacct tctcctggat tcctcatctt ccttctgtga cctgtgtctc 180
catcagtttc tcctccggca tctttttctt acaggattct tacctcagga atagatggac 240
atggcctggc agatgatgca gctgctgctt ctggctttgg tgactgctgc ggggagtgcc 300
cagcccagga gtgcgcgggc caggacggac ctgctcaatg tctgcatgaa cgccaagcac 360
cacaagacac agcccagccc cgaggacgag ctgtatggcc agtgcagtcc ctggaagaag 420
aatgcctgct gcacggccag caccagccag gagctgcaca aggacacctc ccgcctgtac 480
aactttaact gggatcactg tggtcaacca gagctggcgc aaagagcgca ttctgaacgt 540
gcccctgtgc aaagaggact gtgagcgctg gtgggaggac tgtcgcacct cctacacctg 600
caaaagcaac tggcacaaag gctggaattg gacctcaggg attaatgagt gtccggccgg 660
ggccctctgc agcacctttg agtcctactt ccccactcca gccgcccttt gtgaaggcct 720
ctggagccac tccttcaagg tcagcaacta tagtcgaggg agctgccgct gcatccagat 780
gtggtttgac tcagcccagg gcaaccccaa tgaggaggtg gccaagttct atgctgcggc 840
catgaatgct ggggccccgt ctcgtgggat tattgattcc tgatccaaga agggtcctct 900
ggggttcttc caacaaccta ttctaataga caaatccaca tgaaaa 946

<210> 93
<211> 2737
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510733CB1

<400> 93
agcgaacagg gaatgacagt tccaccagaa gacgattaag ccacagcctc taattggaac 60
ggcatttgta cagtcagaga ctcttaccag acatctccag gaatctgtga gccattgtca 120
aaacgtccat tttcatctgg ctgtgaaagt gaggaccaca acaggtaggt attggtagaa 180
acaggagtcc tcagagaagc cccaagatgc agcctgaggg agcagaaaag ggaaaaagct 240
tcaagcagag actggtcttg aagagcagct tagcgaaaga aaccctctct gagttcttgg 300
gcacgttcat cttgattgtc cttggatgtg gctgtgttgc ccaagctatt ctcagtcgag 360
gacgttttgg aggggtcatc actatcaatg ttggattttc aatggcagtt gcaatggcca 420
tttatgtggc tggcggtgtc tctgatggac ttatgtcctt tgctggtgga aaactgctga 480
tcgtgggaga aaatgcaaca gcacacattt ttgcaacata cccagctccg tatctatctc 540
tggcgaacgc atttgcagat caagtggtgg ccaccatgat actcctcata atcgtctttg 600
ccatttttga ctccagaaac ttgggagccc ccagaggcct agagcccatt gccatcggcc 660
tcctgattat tgtcattgct tcctccctgg gactgaacag tggctgtgcc atgaacccag 720
ctcgagacct gagtcccaga cttttcactg ccttggcagg ctgggggttt gaagtcttca 780
gagctggaaa caacttctgg tggattcctg tagtgggccc tttggttggt gctgtcattg 840
gaggcctcat ctatgttctt gtcattgaaa tccaccatcc agagcctgac tcagtcttta 900
aggcagaaca atctgaggac aaaccagaga aatatgaact cagtgtcatc atgtagtggc 960

atgctcagct ctggatttgc agtcagtttg ggattctctt cagaaagatg gcatctaagt 1020
gtctgtgttc ttgtaagcct gaggtggaat ccacccagtt ttgtctgcta gccatatggg 1080
acatctaatt ggaaaagcat ctgcataaaa gtttggaaac aatgaccact tctctaccat 1140
tgtcccccac ccccaccccc cagaataacg ctgactgtcc cctgaaacag ccttctctcc 1200
tgccctgttt atttcatcct cgatgggaat tcttgctagg taagcactaa taactcggca 1260
tcttgacgat agtcccattt gggtggtttc agctgcacta tctgtatgaa atggtgtcac 1320
caaaaccctt ttcttcagta tcgacaaaga ttacattctg agtaccaacc aaaccctaaa 1380
ttgaaagaca aaactatggt ttcagtcaac atattcatga attagggagc taatgggtta 1440
agcttccagt tcccgctatg ctactggatt tgtataaata ctgatattct ccaaacctag 1500
tggtgtaggg agcaagagaa tgcagctgga aggcacaagg ggaggacatt gtggcattca 1560
gaaactgcag gagacaagat gaatttgaga agccaaatgg aatttttaat ggaaaccatt 1620
tatcagatta atctcttgct ctcctgcatt ttagaggaca ccaattaatt tcctggtctt 1680
tagtatataa taacctaaaa taccattgta acctcagtca tgaaaaatac atcactctgt 1740
ctttttagct caaatgtatt ttcctaattg cccacttgag aacagacatt tgacaagtta 1800
tatcaacgac tgtgcttgtc cattatttta cacatgccct agaagccaaa actgaaagcc 1860
actggatcct ggtctagctg aatcttcaga gtgggaggtc tccaaaaaga tattacctta 1920
ttgggcttaa caattcacaa ggcactttca cacccattat ctaatttaat cctcataatg 1980
actatgtgag gcaaatgcca cattgcccat ttttcagata aagaaacaaa atcttaggga 2040
agataagttg agttgtccaa gagcacactg aaagttgaat gttatctaat gcattcctct 2100
acctttcaga agatcagtag ctggctgaga atctttgcca aatcttcctt gctagccaga 2160
agtggaattg gcagcttcta gaatatgtac acctctggac aaaatgttcc tcaatcttaa 2220
gatacaaaga ccctcattgt ctgggtctat tcccacactt actgagtaca gatgaaggaa 2280
agtggtagca atttaatcat aactttcatt tgctgaaaaa cattatgaga aggcctccct 2340
tcctaagcca cctctggtct tgctaagtct tgatcttgct tcctgccagc accaaacatt 2400
acattcaggg gatttcctct ggctcagtct tttccccttg aagttctcta atagatgtta 2460
cttttgacaa aagatcgcct atgagttaca agcaccaggg gatgctctac atcaagggat 2520
gcaccttcag tcaaactgtc aaaaagccca gaattcccaa aggcattagg tttcccaact 2580
gctttgtgct gatatcagaa cagcagaaat taaatgtgaa atgtttctga tgacttatgt 2640
tctacaatct atggacatac gggatttttt ttttcttgct ttgaagctac ctggatattt 2700
cctatttgaa ataaaattgt tcggtcattg ttgaaaa 2737

<210> 94
<211> 2821
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510734CB1

<400> 94
agcgaacagg gaatgacagt tccaccagaa gacgattaag ccacagcctc taattggaac 60
ggcatttgta cagtcagaga ctcttaccag acatctccag gaatctgtga gccattgtca 120
aaacgtccat tttcatctgg ctgtgaaagt gaggaccaca acaggtaggt attggtagaa 180
acaggagtcc tcagagaagc cccaagatgc agcctgaggg agcagaaaag ggaaaaagct 240
tcaagcagag actggtcttg aagagcagct tagcgaaaga aaccctctct gagttcttgg 300
gcacgttcat cttgattgtc cttggatgtg gctgtgttgc ccaagctatt ctcagtcgag 360
gacgttttgg aggggtcatc actatcaatg ttggattttc aatggcagtt gcaatggcca 420
tttatgtggc tggcggtgtc tctggtggtc acatcaaccc agctgtgtct ttagcaatgt 480
gtctctttgg acggatgaaa tggttcaaat tgccatttta tgtgggagcc cagttcttgg 540
gagcctttgt gggggctgca accgtctttg gcatttacta tgatggactt atgtcctttg 600
ctggtggaaa actgctgatc gtgggagaaa atgcaacagc acacattttt gcaacatacc 660
cagctccgta tctatctctg gcgaacgcat ttgcagatca aaacttgggg agcccccaga 720
ggcctagagc ccattgccat cggcctcctg attattgtca ttgcttcctc cctgggactg 780
aacagtggct gtgccatgaa cccagctcga gacctgagtc ccagactttt cactgccttg 840
gcaggctggg ggtttgaagt cttcagagct ggaaacaact tctggtggat tcctgtagtg 900

ggccctttgg ttggtgctgt cattggaggc ctcatctatg ttcttgtcat tgaaatccac 960
catccagagc ctgactcagt ctttaaggca gaacaatctg aggacaaacc agagaaatat 1020
gaactcagtg tcatcatgta gtggcatgct cagctctgga tttgcagtca gtttgggatt 1080
ctcttcagaa agatggcatc taagtgtctg tgttcttgta agcctgaggt ggaatccacc 1140
cagttttgtc tgctagccat atgggacatc taattggaaa agcatctgca taaaagtttg 1200
gaaacaatga ccacttctct accattgtcc cccacccccca ccccccagaa taacgctgac 1260
tgtcccctga aacagccttc tctcctgccc tgtttatttc atcctcgatg ggaattcttg 1320
ctaggtaagc actaataact cggcatcttg acgatagtcc catttgggtg gtttcagctg 1380
cactatctgt atgaaatggt gtcaccaaaa cccttttctt cagtatcgac aaagattaca 1440
ttctgagtac caaccaaacc ctaaattgaa agacaaaact atggtttcag tcaacatatt 1500
catgaattag ggagctaatg ggttaagctt ccagttcccg ctatgctact ggatttgtat 1560
aaatactgat attctccaaa cctagtggtg tagggagcaa gagaatgcag ctggaaggca 1620
caaggggagg acattgtggc attcagaaac tgcaggagac aagatgaatt tgagaagcca 1680
aatggaattt ttaatggaaa ccatttatca gattaatctc ttgctctcct gcattttaga 1740
ggacaccaat taatttcctg gtctttagta tataataacc taaaatacca ttgtaacctc 1800
agtcatgaaa aatacatcac tctgtctttt tagctcaaat gtattttcct aattgcccac 1860
ttgagaacag acatttgaca agttatatca acgactgtgc ttgtccatta ttttacacat 1920
gccctagaag ccaaaactga aagccactgg atcctggtct agctgaatct tcagagtggg 1980
aggtctccaa aaagatatta ccttattggg cttaacaatt cacaaggcac tttcacaccc 2040
attatctaat ttaatcctca taatgactat gtgaggcaaa tgccacattg cccatttttc 2100
agataaagaa acaaaatctt agggaagata agttgagttg tccaagagca cactgaaagt 2160
tgaatgttat ctaatgcatt cctctacctt tcagaagatc agtagctggc tgagaatctt 2220
tgccaaatct tccttgctag ccagaagtgg aattggcagc ttctagaata tgtacacctc 2280
tggacaaaat gttcctcaat cttaagatac aaagaccctc attgtctggg tctattccca 2340
cacttactga gtacagatga aggaaagtgg tagcaattta atcataactt tcatttgctg 2400
aaaaacatta tgagaaggcc tcccttccta agccacctct ggtcttgcta agtcttgatc 2460
ttgcttcctg ccagcaccaa acattacatt caggggattt cctctggctc agtcttttcc 2520
ccttgaagtt ctctaataga tgttactttt gacaaaagat cgcctatgag ttacaagcac 2580
caggggatgc tctacatcaa gggatgcacc ttcagtcaaa ctgtcaaaaa gcccagaatt 2640
cccaaaggca ttaggtttcc caactgcttt gtgctgatat cagaacagca gaaattaaat 2700
gtgaaatgtt tctgatgact tatgttctac aatctatgga catacgggat tttttttttc 2760
ttgctttgaa gctacctgga tatttcctat ttgaaataaa attgttcggt cattgttgaa 2820
a 2821

<210> 95
<211> 3583
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7503977CB1

<400> 95
ctgggcttcg tgccccgacc tggtgtgtca gtgctggggg gatcgggggg ccccgtcctc 60
cagacctggc tgcaggacct gctgcgtcgt gggctggtgc gggctgccca gagcacagga 120
gcctggattg tcactggggg tctgcacacg ggcatcggcc ggcatgttgg tgtggctgta 180
cgggaccatc agatggccag cactggggc accaaggtgg tggccatggg tgtggccccc 240
tggggtgtgg tccggaatag agacaccctc atcaacccca agggctcgtt ccctgcgagg 300
taccggtggc gcggtgaccc ggaggacggg gtccagtttc ccctggacta caactactcg 360
gccttcttcc tggtggacga cggcacacac ggctgcctgg ggggcgagaa ccgcttccgc 420
ttgcgcctgg agtcctacat ctcacagcag aagacgggcg tgggagggac tggaattgac 480
atccctgtcc tgctcctcct gattgatggt gatgagaaga tgttgacgcg aatagagaac 540
gccacccagg ctcagctccc atgtctcctc gtggctggct caggggggagc tgcggactgc 600
ctggcggaga ccctggaaga cactctggcc ccagggagtg gggagccag gcaaggcgaa 660
gcccgagatc gaatcaggcg tttctttccc aaaggggacc ttgaggtcct gcaggcccag 720

gtggagagga ttatgacccg gaaggagctc ctgacagtct attcttctga ggatgggtct 780
gaggaattcg agaccatagt tttgaaggcc cttgtgaagg tccttccatc tcgaagcttc 840
cctcatggac gccctgctga atgaccggcc tgagttcgtg cgcttgctca tttcccacgg 900
cctcagcctg ggccacttcc tgaccccgat gcgcctggcc caactctaca gcgcggcgcc 960
ctccaactcg ctcatccgca accttttgga ccaggcgtcc cacagcgcag gcaccaaagc 1020
cccagcccta aaggggggag ctgcggagct ccggcccccт gacgtggggc atgtgctgag 1080
gatgctgctg ggaagatgt gcgcgccgag gtacccctcc gggggcgcct gggaccctca 1140
cccaggccag ggcttcgggg agagcatgta tctgctctcg gacaaggcca cctcgccgct 1200
ctcgctggat gctggcctcg ggcaggcccc ctggagcgac ctgcttcttt gggcactgtt 1260
gctgaacagg gcacagatgg ccatgtactt ctgggagatg ggttccaatg cagtttcctc 1320
agctcttggg gcctgtttgc tgctccgggt gatggcacgc ctggagcctg acgctgagga 1380
ggcagcacgg aggaaagacc tggcgttcaa gtttgagggg atgggcgttg acctctttgg 1440
cgagtgctat cgcagcagtg aggtgagggc tgcccgcctc ctcctccgtc gctgcccgct 1500
ctggggggat gccacttgcc tccagctggc catgcaagct gacgcccgtg ccttctttgc 1560
ccaggatggg gtacagtctc tgctgacaca gaagtggtgg ggagatatgg ccagcactac 1620
acccatctgg gccctggttc tcgccttctt ttgccctcca ctcatctaca cccgcctcat 1680
caccttcagg aaatcagaag aggagcccac acgggaggag ctagagtttg acatggatag 1740
tgtcattaat ggggaagggc ctgtcgggac ggcggaccca gccgagaaga cgccgctggg 1800
ggtcccgcgc cagtcgggcc gtccgggttg ctgcgggggc cgctgcgggg ggcgccggtg 1860
cctacgccgc tggttccact tctggggcgc gccggtgacc atcttcatgg gcaacgtggt 1920
cagctacctg ctgttcctgc tgcttttctc gcgggtgctg ctcgtggatt tccagccggc 1980
gccgcccggc tccctggagc tgctgctcta tttctgggct ttcacgctgc tgtgcgagga 2040
actgcgccag ggcctgagcg gaggcggggg cagcctcgcc agcgggggcc ccgggcctgg 2100
ccatgcctca ctgagccagc gcctgcgcct ctacctcgcc gacagctgga accagtgcga 2160
cctagtggct ctcacctgct tcctcctggg cgtgggctgc cggctgaccc cgggtttgta 2220
ccacctgggc cgcactgtcc tctgcatcga cttcatggtt ttcacggtgc ggctgcttca 2280
catcttcacg gtcaacaaac agctggggcc caagatcgtc atcgtgagca agatgatgaa 2340
ggacgtgttc ttcttcctct tcttcctcgg cgtgtggctg gtagcctatg gcgtggccac 2400
ggaggggctc ctgaggccac gggacagtga cttcccaagt atcctgcgcc gcgtcttcta 2460
ccgtccctac ctgcagatct tcgggcagat tccccaggag gacatggacg tggccctcat 2520
ggagcacagc aactgctcgt cggagcccgg cttctgggca caccctcctg ggcccaggc 2580
gggcacctgc gtctcccagt atgccaactg gctggtggtg ctgctcctcg tcatcttcct 2640
gctcgtggcc aacatcctgc tggtcaactt gctcattgcc atgttcagtt acacattcgg 2700
caaagtacag ggcaacagcg atctctactg gaaggcgcag cgttaccgcc tcatccggga 2760
attccactct cggcccgcgc tggccccgcc ctttatcgtc atctcccact tgcgcctcct 2820
gctcaggcaa ttgtgcaggc gaccccggag cccccagccg tcctccccgg ccctcgagca 2880
tttccgggtt tacctttcta aggaagccga gcggaagctg ctaacgtggg aatcggtgca 2940
taaggagaac tttctgctgg cacgcgctag ggacaagcgg gagagcgact ccgagcgtct 3000
gaagcgcacg tcccagaagg tggacttggc actgaaacag ctgggacaca tccgcgagta 3060
cgaacagcgc ctgaaagtgc tggagcggga ggtccagcag tgtagccgcg tcctggggtg 3120
ggtggccgag gccctgagcc gctctgcctt gctgccccca ggtgggccgc cacccccctga 3180
cctgcctggg tccaaagact gagccctgct ggcggacttc aaggagaagc ccccacaggg 3240
gattttgctc ctagagtaag gctcatctgg gcctcggccc ccgcacctgg tggccttgtc 3300
cttgaggtga gccccatgtc catctgggcc actgtcagga ccacctttgg gagtgtcatc 3360
cttacaaacc acagcatgcc cggctcctcc cagaaccagt cccagcctgg gaggatcaag 3420
gcctggatcc cgggccgtta tccatctgga ggctgcaggg tccttggggt aacagggacc 3480
acagacccct caccactcac agattcctca cactggggaa ataaagccat ttcagaggaa 3540
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aag 3583

<210> 96
<211> 2125
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7505084CB1

<400> 96
gagaacttta cgcctggatc tcatctaact gacacagaaa ccctgtaagg atccagaggt 60
ctcgttcagg accatggaga gcggcaccag cagccctcag cctccacagt tagatcccct 120
ggatgcgttt ccccagaagg gcttggagcc tggggacatc gcggtgctag ttctgtactt 180
cctctttgtc ctggctgttg gactatggtc cacagtgaag accaaaagag acacagtgaa 240
aggctacttc ctggctggag gggacatggt gtggtggcca gtgggtgcat ccttgtttgc 300
cagcaatgtt ggaagtggac atttcattgg cctggcaggg tcaggtgctg ctacgggcat 360
ttctgtatca gcttatgaac ttaatggctt gttttctgtg ctgatgttgg cctggatctt 420
cctacccatc tacattgctg gtcaggtcac cacgatgcca gaatacctac ggaagcgctt 480
cggtggcatc agaatcccca tcatcctggc tgtactctac ctatttatct acatcttcac 540
caagatctcg gtagacatgt atgcaggtgc catcttcatc cagcagtctt tgcacctgga 600
tctgtacctg gccatagttg ggctactggc catcactgct gtatacacgg ttgctggtgg 660
cctggctgct gtgatctaca cggatgccct gcagacgctg atcatgctta taggagcgct 720
caccttgatg ggctacagtt tcgccgcggt tggtgggatg gaaggactga aggagaagta 780
cttcttggcc ctggctagca accggagtga gaacagcagc tgcgggctgc cccgggaaga 840
tgccttccat attttccgag atccgctgac atctgatctc ccgtggccgg gggtcctatt 900
tggaatgtcc atcccatccc tctggtactg gtgcacggat caggtgattg tccagcggac 960
tctggctgcc aagaacctgt cccatgccaa aggaggtgct ctgatggctg catacctgaa 1020
ggtgctgccc ctcttcataa tggtgttccc tgggatggtc agccgcatcc tcttcccaga 1080
tcaagtggcc tgtgcagatc cagagatctg ccagaagatc tgcagcaacc cctcaggctg 1140
ttcggacatc gcgtatccca aactcgtgct ggaactcctg cccacagtgc cagcaccatc 1200
ttcaccatgg acctctggaa tcacctccgg cctcgggcat ctgagaagga gctcatgatt 1260
gtgggcaggg tgtttgtgct gctgctggtc ctggtctcca tcctctggat ccctgtggtc 1320
caggccagcc agggcggcca gctcttcatc tatatccagt ccatcagctc ctacctgcag 1380
ccgcctgtgg cggtggtctt catcatggga tgtttctgga agaggaccaa tgaaaagggt 1440
gccttctggg gcctgatctc gggcctgctc ctgggcttgg ttaggctggt cctggacttt 1500
atttacgtgc agcctcgatg cgaccagcca gatgagcgcc cggtcctggt gaagagcatt 1560
cactacctct acttctccat gatcctgtcc acggtcaccc tcatcactgt ctccaccgtg 1620
agctggttca cagagccacc ctccaaggag atggtcagcc acctgacctg gtttactcgt 1680
cacgaccccg tggtccagaa ggaacaagca ccaccagcag ctcccttgtc tcttaccctc 1740
tctcagaacg ggatgccaga ggccagcagc agcagcagcg tccagttcga gatggttcaa 1800
gaaaacacgt ctaaaaccca cagctgtgac atgaccccaa agcagtccaa agtggtgaag 1860
gccatcctgt ggctctgtgg aatacaggag aagggcaagg aagagctccc ggccagagca 1920
gaagccatca tagtttccct ggaagaaaac cccttggtga agaccctcct ggacgtcaac 1980
ctcattttct gcgtgagctg cgccatcttt atctggggct atttgcttta gtgtggggtg 2040
aacccagggg tccaaactct gtttctcttc agtgctccat ttttttaatg aaagaaaaaa 2100
taataaagct tttgtttacc aaaaa 2125

<210> 97
<211> 1517
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506950CB1

<400> 97
gaagaaaaca ccgaattctg cttgccgttt cagagcggcg gtgatgaaga caaaattgaa 60
catctacaac atgcagttcc tgctttttgt tttcttggtg tgggaccctg ccaggttggt 120
gctggctaac atccaagaag atgaggctaa aaataacatt accatcttta cgagaattct 180
tgacagactt ctggatggtt acgataatcg gcttagacca ggactgggag acagtattac 240
tgaagtcttc actaacatct acgtgaccag ttttggccct gtctcagata cagatatgga 300
atatacaatt gatgttttct ttcgacaaaa atggaaagat gaacgtttaa aatttaaagg 360

tcctatgaat atccttcgac taaacaattt aatggctagc aaaatctgga ctccagatac 420
ctttttttcac aatgggaaaa aatcagtagc tcataatatg acaatgccaa ataagttgct 480
tcgaattcag gatgatggga ctctgctgta taccatgagg agtaacaact gtcctaacaa 540
tgacaactct aagcatcagt gctcggaatt ctctccccaa agtggcttat gcaactgcca 600
tggactggtt tattgctgtt tgttatgcat ttgtgttctc tgccctaatt gaatttgcaa 660
ctgttaatta cttcaccaaa agaggatggg cttgggatgg gaagagtgta gtaaatgaca 720
agaaaaaaga aaaggcttcc gttatgatac agaacaacgc ttatgcagtg gctgttgcca 780
attatgcccc gaatctttca aaagatccag ttctctccac catctccaag agtgcaacca 840
cgccagaacc caacaagaag ccagaaaaca agccagctga agcaaagaaa actttcaaca 900
gtgttagcaa aattgacaga atgtccagaa tagtttttcc agttttgttt ggtaccttta 960
atttagttta ctgggctaca tatttaaaca gagaacctgt attaggggtc agtccttgaa 1020
ttgagaccca tgttatcttt gggatgtata gcaacattaa atttggtttg ttttgctatg 1080
tacagtctga ctaataactg ctaatttgtg atccaacatg tacagtatgt atatagtgac 1140
atagcttacc agtagacctt taatggagac atgcatttgc taactcatgg aactgcagac 1200
agaaagcact ccatgcgaaa acagccattg cctttttttaa agatttaccc taggacctga 1260
tttaaagtga atttcaaatg acctgattaa tttcctattc ttccaaatga gatgaaaatg 1320
gggatcctgt acaacccttt gtggacactt ttggtttagc tcttaagtag ggtattttc 1380
tactgttgcc taatatgatg gaagtaacat tgtcattcta gatgaatctt tgaagtacca 1440
acattgttct ggaatcagct cggtccagat gctcaaggtc cctgtaatgt attggaagct 1500
ggtaccctaa gaaaaca 1517

<210> 98
<211> 1694
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506951CB1

<400> 98
gaagaaaaca ccgaattctg cttgccgttt cagagcggcg gtgatgaaga caaaattgaa 60
catctacaac atgcagttcc tgctttttgt tttcttggtg tgggaccctg ccaggttggt 120
gctggctaac atccaagaag atgaggctaa aaataacatt accatcttta cgagaattct 180
tgacagactt ctggatggtt acgataatcg gcttagacca ggactgggag acagtattac 240
tgaagtcttc actaacatct acgtgaccag ttttggccct gtctcagata cagatatgga 300
atatacaatt gatgttttct ttcgacaaaa atggaaagat gaacgtttaa aatttaaagg 360
tcctatgaat atccttcgac taaacaattt aatggctagc aaaatctgga ctccagatac 420
ctttttttcac aatgggaaga aatcagtagc tcataatatg acaatgccaa ataagttgct 480
tcgaattcag gatgatggga ctctgctgta taccatgagg cttacagttc aagctgaatg 540
cccaatgcac ttggaggatt tcccaatgga tgctcattca tgtcctctga aatttggcag 600
ctatgcatat acaacttcag aggtcactta tatttggact tacaatgcat ctgattcagt 660
acaggttgct cctgatggct ctaggttaaa tcaatatgac ctgctgggcc aatcaatcgg 720
aaaggagaca attaaatcca gtacaggtga atatactgta atgacagctc atttccacct 780
gaaaagaaaa attgggtatt ttgtgattca aacctatctg ccttgcatca tgactgtcat 840
tctctcccaa gtttcattct ggcttaacag agaatctgtg cctgcaagaa ctgtgtttga 900
aaaaagaaaa ggcttccgtt atgatacaga acaacgctta tgcagtggct gttgccaatt 960
atgccccgaa tctttcaaaa gatccagttc tctccaccat ctccaagagt gcaaccacgc 1020
cagaacccaa caagaagcca gaaaacaagc cagctgaagc aaagaaaact ttcaacagtg 1080
ttagcaaaat tgacagaatg tccagaatag tttttccagt tttgtttggt acctttaatt 1140
tagtttactg ggctacatat ttaaacagag aacctgtatt aggggtcagt ccttgaattg 1200
agacccatgt tatctttggg atgtatagca acattaaatt tggtttgttt tgctatgtac 1260
agtctgacta ataactgcta atttgtgatc caacatgtac agtatgtata tagtgacata 1320
gcttaccagt agacctttaa tggagacatg catttgctaa ctcatggaac tgcagacaga 1380
aagcactcca tgcgaaaaca gccattgcct tttttaaaga tttaccctag gacctgattt 1440
aaagtgaatt tcaaatgacc tgattaattt cctattcttc caaatgagat gaaaatgggg 1500

atcctgtaca accctttgtg gacacttttg gtttagctct taagtagggg tattttctac 1560
tgttgcctaa tatgatggaa gtaacattgt cattctagat gaatctttga agtaccaaca 1620
ttgttctgga atcagctcgg tccagatgct caaggtccct gtaatgtatt ggaagctggt 1680
accctaagaa aaca 1694

<210> 99
<211> 1102
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506954CB1

<400> 99
ctctctctct ctctctctct ctctctctct ctctctctct ctcccaagtt tcctatctcg 60
tcaagatcag ggcaaaagaa gaaaacaccg aattctgctt gccgtttcag agcggcggtg 120
atgaagacaa aattgaacat ctacaacatg cagttcctgc tttttgtttt cttggtgtgg 180
gaccctgcca ggttggtgct ggctaacatc caagaagatg aggctaaaaa taacattacc 240
atctttacga gaattcttga cagacttctg gatggttacg ataatcggct tagaccagga 300
ctgggagaaa aaagaaaagg cttccgttat gatacagaac aacgcttatg cagtggctgt 360
tgccaattat gccccgaatc tttcaaaaga tccagttctc tccaccatct ccaagagtgc 420
aaccacgcca gaacccaaca agaagccaga aaacaagcca gctgaagcaa agaaaacttt 480
caacagtgtt agcaaaattg acagaatgtc cagaatagtt tttccagttt tgtttggtac 540
ctttaattta gtttactggg ctacatattt aaacagagaa cctgtattag ggtcagtcc 600
ttgaattgag acccatgtta tctttgggat gtatagcaac attaaatttg gtttgttttg 660
ctatgtacag tctgactaat aactgctaat ttgtgatcca acatgtacag tatgtatata 720
gtgacatagc ttaccagtag acctttaatg gagacatgca tttgctaact catggaactg 780
cagacagaaa gcactccatg cgaaaacagc cattgccttt tttaaagatt taccctagga 840
cctgatttaa agtgaatttc aaatgacctg attaatttcc tattcttcca aatgagatga 900
aaatggggat cctgtacaac cctttgtgga cacttttggt ttagctctta agtaggggta 960
ttttctactg ttgcctaata tgatggaagt aacattgtca ttctagatga atctttgaag 1020
taccaacatt gttctggaat cagctcggtc cagatgctca aggtccctgt aatgtattgg 1080
aagctggtac cctaagaaaa ca 1102

<210> 100
<211> 1744
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506956CB1

<400> 100
gaagaaaaca ccgaattctg cttgccgttt cagagcggcg gtgatgaaga caaaattgaa 60
catctacaac atgcagttcc tgctttttgt ttcttggtg tgggaccctg ccaggttggt 120
gctggctaac atccaagaag atgaggctaa aaataacatt accatcttta cgagaattct 180
tgacagactt ctggatggtt acgataatcg gcttagacca ggactgggag acagtattac 240
tgaagtcttc actaacatct acgtgaccag ttttggccct gtctcagata cagatatgga 300
atatacaatt gatgttttct ttcgacaaaa atggaaagat gaacgtttaa aatttaaagg 360
tcctatgaat atccttcgac taaacaattt aatggctagc aaaatctgga ctccagatac 420
ctttttttcac aatgggaaga aatcagtagc tcataatatg acaatgccaa ataagttgct 480
tcgaattcag gatgatggga ctctgctgta taccatgagg cttacagttc aagctgaatg 540
cccaatgcac ttggaggatt tcccaatgga tgctcattca tgtcctctga aatttggcag 600
ctatgcatat acaacttcag aggtcactta tatttggact tacaatgcat ctgattcagt 660

acaggttgct cctgatggct ctaggttaaa tcaatatgac ctgctgggcc aatcaatcgg 720
aaaggagaca attaaatcca gtacaggagt aacaactgtc ctaacaatga caactctaag 780
catcagtgct cggaattctc tccccaaagt ggcttatgca actgccatgg actggtttat 840
tgctgtttgt tatgcatttg tgttctctgc cctaattgaa tttgcaactg ttaattactt 900
caccaaaaga ggatgggctt gggatgggaa gagtgtagta aatgacaaga aaaaagaaaa 960
ggcttccgtt atgatacaga acaacgctta tgcagtggct gttgccaatt atgccccgaa 1020
tctttcaaaa gatccagttc tctccaccat ctccaagagt gcaaccacgc cagaacccaa 1080
caagaagcca gaaaacaagc cagctgaagc aaagaaaact ttcaacagtg ttagcaaaat 1140
tgacagaatg tccagaatag tttttccagt tttgtttggt acctttaatt tagtttactg 1200
ggctacatat ttaaacagag aacctgtatt aggggtcagt ccttgaattg agacccatgt 1260
tatctttggg atgtatagca acattaaatt tggtttgttt tgctatgtac agtctgacta 1320
ataactgcta atttgtgatc caacatgtac agtatgtata tagtgacata gcttaccagt 1380
agacctttaa tggagacatg catttgctaa ctcatggaac tgcagacaga aagcactcca 1440
tgcgaaaaca gccattgcct tttttaaaga tttaccctag gacctgattt aaagtgaatt 1500
tcaaatgacc tgattaattt cctattcttc caaatgagat gaaaatgggg atcctgtaca 1560
accctttgtg gacacttttg gtttagctct taagtagggg tattttctac tgttgcctaa 1620
tatgatggaa gtaacattgt cattctagat gaatctttga agtaccaaca ttgttctgga 1680
atcagctcgg tccagatgct caaggtccct gtaatgtatt ggaagctggt accctaagaa 1740
aaca 1744

<210> 101
<211> 1753
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506959CB1

<400> 101
gaagaaaaca ccgaattctg cttgccgttt cagagcggcg gtgatgaaga caaaattgaa 60
catctacaac atgcagttcc tgctttttgt tttcttggtg tgggaccctg ccaggttggt 120
gctggctaac atccaagaag atgaggctaa aaataacatt accatcttta cgagaattct 180
tgacagactt ctggatggtt acgataatcg gcttagacca ggactgggag acagtattac 240
tgaagtcttc actaacatct acgtgaccag ttttggccct gtctcagata cagatatgga 300
atatacaatt gatgttttct ttcgacaaaa atggaaagat gaacgtttaa aatttaaagg 360
tcctatgaat atccttcgac taaacaattt aatggctagc aaaatctgga ctccagatac 420
ctttttttcac aatgggaaga aatcagtagc tcataatatg acaatgccaa ataagttgct 480
tcgaattcag gatgatggga ctctgctgta taccatgagg cttacagttc aagctgaatg 540
cccaatgcac ttggaggatt tcccaatgga tgctcattca tgtcctctga aatttggcag 600
ctgtgaatat actgtaatga cagctcattt ccacctgaaa agaaaaattg ggtattttgt 660
gattcaaacc tatctgcctt gcatcatgac tgtcattctc tcccaagttt cattctggct 720
taacagagaa tctgtgcctg caagaactgt gtttggagta acaactgtcc taacaatgac 780
aactctaagc atcagtgctc ggaattctct ccccaaagtg gcttatgcaa ctgccatgga 840
ctggtttatt gctgtttgtt atgcatttgt gttctctgcc ctaattgaat ttgcaactgt 900
taattacttc accaaaagag gatgggcttg ggatgggaag agtgtagtaa atgacaagaa 960
aaaagaaaag gcttccgtta tgatacagaa caacgcttat gcagtggctg ttgccaatta 1020
tgccccgaat ctttcaaaag atccagttct ctccaccatc tccaagagtg caaccacgcc 1080
agaacccaac aagaagccag aaaacaagcc agctgaagca aagaaaactt tcaacagtgt 1140
tagcaaaatt gacagaatgt ccagaatagt ttttccagtt ttgtttggta cctttaattt 1200
agtttactgg gctacatatt taaacagaga acctgtatta ggggtcagtc cttgaattga 1260
gacccatgtt atctttggga tgtatagcaa cattaaattt ggtttgtttt gctatgtaca 1320
gtctgactaa taactgctaa tttgtgatcc aacatgtaca gtatgtatat agtgacatag 1380
cttaccagta gacctttaat ggagacatgc atttgctaac tcatggaact gcagacagaa 1440
agcactccat gcgaaaacag ccattgcctt ttttaaagat ttaccctagg acctgattta 1500
aagtgaattt caaatgacct gattaatttc ctattcttcc aaatgagatg aaaatgggga 1560

tcctgtacaa ccctttgtgg acacttttgg tttagctctt aagtaggggt attttctact 1620
gttgcctaat atgatggaag taacattgtc attctagatg aatctttgaa gtaccaacat 1680
tgttctggaa tcagctcggt ccagatgctc aaggtccctg taatgtattg gaagctggta 1740
ccctaagaaa aca 1753

<210> 102
<211> 1609
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506960CB1

<400> 102
ctctctctct ctctctctct ctctctctct ctctctctct ctcccaagtt tcctatctcg 60
tcaagatcag ggcaaaagaa gaaaacaccg aattctgctt gccgtttcag agcggcggtg 120
atgaagacaa aattgaacat ctacaacatg cagttcctgc tttttgtttt cttggtgtgg 180
gaccctgcca ggttggtgct ggctaacatc caagaagatg aggctaaaaa taacattacc 240
atctttacga gaattcttga cagacttctg gatggttacg ataatcggct tagaccagga 300
ctgggaggaa tatacaattg atgttttctt tcgacaaaaa tggaaagatg aacgtttaaa 360
atttaaaggt cctatgaata tccttcgact aaacaattta atggctagca aaatctggac 420
tccagatacc ttttttcaca atgggaagaa atcagtagct cataatatga caatgccaaa 480
taagttgctt cgaattcagg atgatgggac tctgctgtat accatgaggc ttacagttca 540
agctgaatgc ccaatgcact tggaggattt cccaatggat gctcattcat gtcctctgaa 600
atttggcagc tgagtaacaa ctgtcctaac aatgacaact ctaagcatca gtgctcggaa 660
ttctctcccc aaagtggctt atgcaactgc catggactgg tttattgctg tttgttatgc 720
atttgtgttc tctgccctaa ttgaatttgc aactgttaat tacttcacca aaagaggatg 780
ggcttgggat gggaagagtg tagtaaatga caagaaaaaa gaaaaggctt ccgttatgat 840
acagaacaac gcttatgcag tggctgttgc caattatgcc ccgaatcttt caaaagatcc 900
agttctctcc accatctcca agagtgcaac cacgccagaa cccaacaaga agccagaaaa 960
caagccagct gaagcaaaga aaactttcaa cagtgttagc aaaattgaca gaatgtccag 1020
aatagttttt ccagttttgt ttggtacctt taatttagtt tactgggcta catatttaaa 1080
cagagaacct gtattagggg tcagtccttg aattgagacc catgttatct ttgggatgta 1140
tagcaacatt aaatttggtt tgttttgcta tgtacagtct gactaataac tgctaatttg 1200
tgatccaaca tgtacagtat gtatatagtg acatagctta ccagtagacc tttaatggag 1260
acatgcattt gctaactcat ggaactgcag acagaaagca ctccatgcga aaacagccat 1320
tgcctttttt aaagatttac cctaggacct gatttaaagt gaatttcaaa tgacctgatt 1380
aatttcctat tcttccaaat gagatgaaaa tggggatcct gtacaaccct ttgtggacac 1440
ttttggttta gctcttaagt aggggtattt tctactgttg cctaatatga tggaagtaac 1500
attgtcattc tagatgaatc tttgaagtac caacattgtt ctggaatcag ctcggtccag 1560
atgctcaagg tccctgtaat gtattggaag ctggtaccct aagaaaaca 1609

<210> 103
<211> 1930
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510540CB1

<400> 103
attgacagga ctcccaacta gtacaatgac agaagataag gtcactggga ccctggtttt 60
cactgtcatc actgctgtgc tgggttcctt ccagtttgga tatgacattg gtgtgatcaa 120
tgcacctcaa cagaatcaaa gccatgttag tagcaaacat tctgtcatta gttggagctc 180

tcttgatggg gttttcaaaa ttgggaccat ctcatatact tataattgct ggaagaagca 240
tatcaggact atattgtggg ctaatttcag gcctggttcc tatgtatatc ggtgaaattg 300
ctccaaccgc tctcagggga gcacttggca cttttcatca gctggccatc gtcacgggca 360
ttcttattag tcagattatt ggtcttgaat ttatcttggg caattatgat ctgtggcaca 420
tcctgcttgg cctgtctggt gtgcgagcca tccttcagtc tctgctactc tttttctgtc 480
cagaaagccc cagatacctt tacatcaagt tagatgagga agtcaaagca aaacaaagct 540
tgaaaagact cagaggatat gatgatgtca ccaaagatat taatgaaatg agaaaagaaa 600
gagaagaagc atcgagtgag cagaaagtct ctataattca gctcttcacc aattccagct 660
accgacagcc tattctagtg gcactgatgc tgcatgtggc tcagcaattt tccggaatca 720
atggcatttt ttactactca accagcattt ttcagacggc tggtatcagc aaacctgttt 780
atgcaaccat tggagttggc gctgtaaaca tggttttcac tgctgtctct gtattccttg 840
tggagaaggc agggcgacgt tctctctttc taattggaat gagtgggatg tttgtttgtg 900
ccatcttcat gtcagtggga cttgtgctgc tgaataagtt ctcttggatg agttatgtga 960
gcatgatagc catcttcctc tttgtcagct tctttgaaat tgggccaggc ccgatcccct 1020
ggttcatggt ggctgagttt ttcagtcaag gaccacgtcc tgctgcttta gcaatagctg 1080
cattcagcaa ttggacctgc aatttcattg tagctctgtg tttccagtac attgcggact 1140
tctgtggacc ttatgtgttt ttcctctttg ctggagtgct cctggccttt accctgttta 1200
catttttaa agttccagaa accaaaggaa agtcttttga ggaaattgct gcagaattcc 1260
aaaagaagag tggctcagcc cacaggccaa aagctgctgt agaaatgaaa ttcctaggag 1320
ctacagagac tgtgtaaaaa aaaaaccctg ctttttgaca tgaacagaaa caataaggga 1380
accgtctgtt tttaaatgat gattccttga gcattttata tccacatctt taagtattgt 1440
tttattttta tgtgctctca tcagaaatgt catcaaatat taccaaaaaa gtattttttt 1500
aagttagaga atatattttt gatggtaaga ctgtaattaa gtaaaccaaa aaggctagtt 1560
tattttgtta cactaaaggg caggtggttc taatattttt agctctgttc tttataacaa 1620
ggttcttcta aaattgaaga gatttcaaca tatcattttt ttaacacata actagaaacc 1680
tgaggatgca acaaatattt atatatttga atatcattaa attggaattt tcttacccat 1740
atatcttatg ttaaaggaga tatggctagt ggcaataagt tccatgttaa aatagacaac 1800
tcttccattt attgcactca gctttttttct tgagtactag aatttgtatt ttgcttaaaa 1860
ttttactttt gttctgtatt ttcatgtgga atggattata gagtatacta aaaaatgtct 1920
atagagaaaa 1930

<210> 104
<211> 1205
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510545CB1

<400> 104
cttggcggag gcgaggaggc cgcggggacg ggaggcgagg ccggccgggc ccccgaagcc 60
atggagaacg cgcacaccaa gacggtggag gaggtgctgg gccacttcgg cgtcaacgag 120
agtgaatctg tctctgtcat caagcacact gatcccgtcc ctgacccacg agctgtcaac 180
caagataaaa agaacatgct gttttctgtg gccctggctg tagcagccat tcctgaaggt 240
ctgcctgcag tcatcaccac ctgcctggct cttggaactc gcagaatggc aaagaaaaat 300
gccattgttc gaagcctccc gtctgtggaa acccttggtt gtacttctgt tatctgctca 360
gacaagactg gtacacttac aacaaaccag atgtcagtct gcaggatgtt cattctggac 420
agagtggaag atcacaccgc tgaacgtgac ccagtggctg atggtgctga aaatctcctt 480
gcccgtgatt ctcatggatg agacgctcaa gtttgtggcc cgcaactacc tggaacctgg 540
taaagagtgt gtgcagcctg ccaccaaatc ctgctcgttc tcggcatgca ccgatgggat 600
ttcctggccg tttgtgctgc tcataatgcc cctggtgatc tgggtctata gcacagacac 660
taactttagc gatatgttct ggtcttgact gacagttttc cataaagaag atgtttaact 720
taatcaatta attttttat tgtttaaagc aactgtctat ttctgctgaa ttttcacatg 780
aacatactgg ctggtgatgg aggtttcata ctctagattt tgttttgctt tttctgactc 840
cagtggggca agattttcct tttttataca cataattaaa gtgtccattg acatgtacag 900

agaactaaca ctattttatg caaatatttt tttgtagatg aaaaagcatg tacagtgttc 960
tgtttaatac tcatccttgt ataaaaaaaa tagttgagcc agcagacatt gtcagcaaat 1020
taattggcag cagattttag gaaatgaatg tgtgtggttt tttttctaaa actaaatagc 1080
atgtattgtg tcttttgcat gatgatccgg atttaatttg atatcacagt ctaattttta 1140
ttcataagcc attttcgcac ggcagagtct gcactcatca gatgtttggt gcactcgcac 1200
cactg 1205

<210> 105
<211> 1790
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510654CB1

<400> 105
cggcggttca ggcgccagag ctggccgatc ggcgttggcc gccgacatga cgcccgagga 60
cccagaggaa acccagccgc ttctggggcc tcctggcggc agcgcgcccc gcggccgccg 120
cgtcttcctc gccgccttcg ccgctgccct gggcccactc agcttcggct tcgcgctcgg 180
ctacagctcc ccggccatcc ctagcctgca gcgcgccgcg cccccggccc cgcgcctgga 240
cgacgccgcc gcctcctggt tcggggctgt cgtgaccctg ggtgccgcgg cgggggagt 300
gctgggcggc tggctggtgg accgcgccgg gcgcaagctg agcctcttgc tgtgctccgt 360
gcccttcgtg gccggctttg ccgtcatcac cgcggcccag gacgtgtgga tgctgctggg 420
gggccgcctc ctcaccggcc tggcctgcgg tgttgcctcc ctagtggccc cggtctacat 480
ctccgaaatc gcctacccag cagtccgggg gttgctcggc tcctgtgtgc agctaatggt 540
cgtcgtcggc atcctcctgg cctacctggc aggctgggtg ctggagtggc gctggctggc 600
tgtgctgggc tgcgtgcccc cctccctcat gctgcttctc atgtgcttca tgcccgagac 660
cccgcgcttc ctgctgactc agcacaggcg ccaggaggcc atggccgccc tgcggttcct 720
gtggggctcc gagcagggct gggaagaccc ccccatcggg gctgagcaga gctttcacct 780
ggccctgctg cggcagcccg gcatctacaa gcccttcatc atcggcgtct ccctgatggc 840
cttccagcag ctgtcggggg tcaacgccgt catgttctat gcagagacca tctttgaaga 900
ggccaagttc aaggacagca gcctggcctc ggtcgtcgtg ggtgtcatcc aggtgctgtt 960
cacagctgtg gcggctctca tcatggacag agcagggcgg aggctgctcc tggtcttgtc 1020
aggaggtcct caggccctat ggagccttct ggcttgcctc cgctttctgc atcttcagtg 1080
tccttttcac tttgttctgt gtccctgaaa ctaaaggaaa gactctggaa caaatcacag 1140
cccattttga ggggcgatga cagccactca ctaggggatg gagcaagcct gtgactccaa 1200
gctgggccca agcccagagc ccctgcctgc cccaggggag ccagaatcca gccccttgga 1260
gccttggtct gcagggtccc tccttcctgt catgctccct ccagcccatg acccggggct 1320
aggaggctca ctgcctcctg ttccagctcc tgctgctgct ctgaggactc aggaacacct 1380
tcgagctttg cagacctgcg gtcagccctc catgcgcaag actaaagcag cggaagagga 1440
ggtgggcctc taggatcttt gtcttctggc tggaggtgct tttggaggtt gggtgctggg 1500
cattcagtcg ctcctctcac gcggctgcct tatcgggaag gaaatttgtt tgccaaataa 1560
agactgacac agaaaatcag gtcagtgtct ctgggctttg tgcaagctca gtttgaaaag 1620
ggtttattcc catcactgcc caggacaccc tgtggcttta cttgctcatg gtcagccaag 1680
cttacccttc acactgagaa gtcatttctg gctacttcct tgggctcagt tccctgggtc 1740
atcagccatc aaatcttgtt gagtttaaaa aataaacagc aaaaaaaaaa 1790

<210> 106
<211> 3824
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510660CB1

<400> 106
gccagcctgt tctgttgccc tggctcttcc tagtccaggc tgccatggcg gcgctcaggg 60
cttaccggaa gtaaaacttc ggaagtgagg cgttcctctg cccggaagtg agcgcggcgc 120
taggaaagat ggcggcagcg gcggcggtgg gcaacgcggt gccctgcggg gcccggcctt 180
gcggggtccg gcctgacggg cagcccaagc ccgggccgca gccgcgcgcg ctccttgccg 240
ccgggccggc gctcatagcg aacggtgacg agctggtggc tgccgtgtgg ccgtaccggc 300
ggttggcgct gttgcggcgc ctcacggtgc tgccattcgc cgggctgctt tacccggcct 360
ggttgggtgc cgcagccgct ggctgctggg gctggggcag cagttgggtg cagatccccg 420
aagctgcgct gctcgtgctt gccaccatct gcctcgcgca cgcgctcact gtcctctcgg 480
ggcattggtc tgtgcacgcg cattgcgcgc tcacctgcac cccggagtac gaccccagca 540
aagcgacctt tgtgaaggtg gtgccaaccc ccaacaatgg ctccacggag ctcgtggccc 600
tgcaccgcaa tgagggcgaa gacgggcttg aggtgctgtc cttcgaattc cagaagatca 660
agtattccta cgatgccctg gagaagaagc agtttctccc cgtggccttt cctgtgggaa 720
acgccttctc atactatcag agcaacagag gcttccagga agactcagag atccgagcag 780
ctgagaagaa atttgggagc aacaaggccg agatggtggt gcctgacttc tcggagcttt 840
tcaaggagag agccacagcc cccttctttg tatttcaggt gttctgtgtg ggctctggt 900
gcctggatga gtactggtac tacagcgtct ttacgctatc catgctggtg gcgttcgagg 960
cctcgctggt gcagcagcag atgcggaaca tgtcggagat ccggaagatg ggcaacaagc 1020
cccacatgat ccaggtctac cgaagccgca agtggaggcc cattgccagt gatgagatcg 1080
taccagggga catcgtctcc atcggccgct ccccacagga gaacctggtg ccatgtgacg 1140
tgcttctgct gcgaggccgc tgcatcgtag acgaggccat gctcacgggg gagtccgtgc 1200
cacagatgaa ggagcccatc gaagacctca gcccagaccg ggtgctggac ctccaggctg 1260
attcccggct gcacgtcatc ttcgggggca ccaaggtggt gcagcacatc cccccacaga 1320
aagccaccac gggcctgaag ccggttgaca gcgggtgcgt ggcctacgtc ctgcggaccg 1380
gattcaacac atcccagggc aagctgctgc gcaccatcct cttcggggtc aagagggtga 1440
ctgcgaacaa cctggagacc ttcatcttca tcctcttcct cctggtgttt gccatcgctg 1500
cagctgccta tgtatggatt gaaggtacca aggaccccag ccggaaccgc tacaagctgt 1560
ttctggagtg caccctgatc ctcacctcgg tcgtgcctcc tgagctgccc atcgagctgt 1620
ccctggccgt caacacctcc ctcatcgccc tggccaagct ctacatgtac tgcacagagc 1680
ccttccggat cccctttgct ggcaaggtcg aggtgtgctg ctttgacaag acggggacgt 1740
tgaccagtga cagcctggtg gtgcgcggtg tggccgggct gagagacggg aaggaggtga 1800
ccccagtgtc cagcatccct gtagaaacac accgggccct ggcctcgtgc cactcgctca 1860
tgcagctgga cgacggcacc ctcgtgggtg accctctaga gaaggccatg ctgacggccg 1920
tggactggac gctgaccaaa gatgagaaag tattcccccg aagtattaaa actcaggggc 1980
tgaaaattca ccagcgcttt cattttgcca gtgccctgaa gcgaatgtcc gtgcttgcct 2040
cgtatgagaa gctgggctcc accgacctct gctacatcgc ggccgtgaag gggcccccg 2100
aaactctgca ctccatgttc tcccagtgcc cgccccgacta ccaccacatc cacaccgaga 2160
tctcccggga aggagcccgc gtcctggcgc tggggtacaa ggagctggga cacctcactc 2220
accagcaggt ggtcatgatc acgggagaca acccgctcac tgcatgccac gtggcccagg 2280
agctgcactt cattgaaaag gcccacacgc tgatcctgca gcctccctcc gagaaaggcc 2340
ggcagtgcga gtggcgctcc attgacggca gcatcgtgct gcccctggcc cggggctccc 2400
caaaggcact ggccctggag tacgcactgt gcctcacagg cgacggcttg gcccacctgc 2460
aggccaccga cccccagcag ctgctccgcc tcatccccca tgtgcaggtg ttcgcccgtg 2520
tggctcccaa gcagaaggag tttgtcatca ccagcctgaa ggagctgggc tacgtgaccc 2580
tcatgtgtgg ggatggcacc aacgacgtgg gcgccctgaa gcatgctgac gtgggtgtgg 2640
cgctcttggc caatgcccct gagcgggttg tcgagcggcg acggcggccc cgggacagcc 2700
caaccctgag caacagtggc atcagagcca cctccaggac agccaagcag cggtcggggc 2760
tccctccctc cgaggagcag ccaacctccc agagggaccg cctgagccag gtgctgcgag 2820
acctcgagga cgagagtacg cccattgtga aactggggga tgccagcatc gcagcaccct 2880
tcacctccaa gctctcatcc atccagtgca tctgccacgt gatcaagcag ggccgctgca 2940
cgctggtgac cacgctacag atgttcaaga tcctggcgct caatgccctc atcctggcct 3000
acagccagag cgtcctctac ctggagggag tcaagttcag tgacttccag gccaccctac 3060
aggggctgct gctggccggc tgcttcctct tcatctcccg ttccaagccc ctcaagaccc 3120
tctcccgaga acggcccctg cccaacatct tcaacctgta caccatcctc accgtcatgc 3180
tccagttctt tgtgcacttc ctgagccttg tctacctgta ccgtgaggcc caggcccgga 3240
gccccgagaa gcaggagcag ttcgtggact tgtacaagga gtttgagcca agcctggtca 3300

acagcaccgt ctacatcatg gccatggcca tgcagatggc caccttcgcc atcaattaca 3360
aaggcccgcc cttcatggag agcctgcccg agaacaagcc cctggtgtgg agtctggcag 3420
tttcactcct ggccatcatt ggcctgctcc tcggctcctc gcccgacttc aacagccagt 3480
ttggcctcgt ggacatccct gtggaggtcc tgctcctgga cttctgcctg gcgctcctgg 3540
ccgaccgcgt cctgcagttc ttcctgggga ccccgaagct gaaagtgcct tcctgagatg 3600
gcagtgctgg tacccactgc ccaccctggc tgccgctggg cgggaacccc aacagggccc 3660
cgggagggaa ccctgccccc aaccccccac agcaaggctg tacagtctcg cccttggaag 3720
actgagctgg gacccccaca gccatccgct ggcttggcca gcagaaccag ccccaagcca 3780
gcacctttgg taaataaagc agcatctgag attttaaaaa aaaa 3824

<210> 107
<211> 3770
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510661CB1

<400> 107
gccagcctgt tctgttgccc tggctcttcc tagtccaggc tgccatggcg gcgctcaggg 60
cttaccggaa gtaaaacttc ggaagtgagg cgttcctctg cccggaagtg agcgcggcgc 120
taggaaagat ggcggcagcg gcggcggtgg gcaacgcggt gccctgcggg gcccggcctt 180
gcggggtccg gcctgacggg cagcccaagc ccgggccgca gccgcgcgcg ctccttgccg 240
ccgggccggc gctcatagcg aacggtgacg agctggtggc tgccgtgtgg ccgtaccggc 300
ggttggcgct gttgcggcgc ctcacggtgc tgccattcgc cgggctgctt tacccggcct 360
ggttgggtgc cgcagccgct ggctgctggg gctggggcag cagttgggtg cagatccccg 420
aagctgcgct gctcgtgctt gccaccatct gcctcgcgca cgcgctcact gtcctctcgg 480
ggcattggtc tgtgcacgcg cattgcgcgc tcacctgcac cccggagtac gaccccagca 540
aagcgacctt tgtgaaggtg gtgccaaccc ccaacaatgg ctccacggag ctcgtggccc 600
tgcaccgcaa tgagggcgaa gacgggcttg aggtgctgtc cttcgaattc cagaagatca 660
agtattccta cgatgccctg gagaagaagc agtttctccc cgtggccttt cctgtgggaa 720
acgccttctc atactatcag agcaacagag gcttccagga agactcagag atccgagcag 780
ctgagaagaa atttgggagc aacaaggccg agatggtggt gcctgacttc tcggagcttt 840
tcaaggagag agccacagcc cccttctttg tatttcaggt gttctgtgtg gggctctggt 900
gcctggatga gtactggtac tacagcgtct ttacgctatc catgctggtg gcgttcgagg 960
cctcgctggt gcagcagcag atgcggaaca tgtcggagat ccggaagatg ggcaacaagc 1020
cccacatgat ccaggtctac cgaagccgca agtggaggcc cattgccagt gatgagatcg 1080
taccagggga catcgtctcc atcggccgct ccccacagga gaacctggtg ccatgtgacg 1140
tgcttctgct gcgaggccgc tgcatcgtag acgaggccat gctcacgggg gagtccgtgc 1200
cacagatgaa ggagcccatc gaagacctca gcccagaccg ggtgctggac ctccaggctg 1260
attcccggct gcacgtcatc ttcgggggca ccaaggtggt gcagcacatc cccccacaga 1320
aagccaccac gggcctgaag ccggttgaca gcgggtgcgt ggcctacgtc ctgcggaccg 1380
gattcaacac atcccagggc aagctgctgc gcaccatcct cttcggggtc aagagggtga 1440
ctgcgaacaa cctggagacc ttcatcttca tcctcttcct cctggtgttt gccatcgctg 1500
cagctgccta tgtatggatt gaaggtacca aggaccccag ccggaaccgc tacaagctgt 1560
ttctggagtg caccctgatc ctcacctcgg tcgtgcctcc tgagctgccc atcgagctgt 1620
ccctggccgt caacacctcc ctcatcgccc tggccaagct ctacatgtac tgcacagagc 1680
ccttccggat ccccttttgct ggcaaggtcg aggtgtgctg ctttgacaag acggggacgt 1740
tgaccagtga cagcctggtg gtgcgcggtg tggccgggct gagagacggg aaggaggtga 1800
ccccagtgtc cagcatccct gtagaaacac accgggccct ggcctcgtgc cactcgctca 1860
tgcagctgga cgacggcacc ctcgtgggtg accctctaga gaaggccatg ctgacggccg 1920
tggactggac gctgaccaaa gtgccctgaa gcgaatgtcc gtgcttgcct cgtatgagaa 1980
gctgggctcc accgacctct gctacatcgc ggccgtgaag gggggccccg aaactctgca 2040
ctccatgttc tcccagtgcc cgcccgacta ccaccacatc cacaccgaga tctcccggga 2100
aggagcccgc gtcctggcgc tggggtacaa ggagctggga cacctcactc accagcaggt 2160

ggtcatgatc acgggagaca acccgctcac tgcatgccac gtggcccagg agctgcactt 2220
cattgaaaag gcccacacgc tgatcctgca gcctccctcc gagaaaggcc ggcagtgcga 2280
gtggcgctcc attgacggca gcatcgtgct gcccctggcc cggggctccc caaaggcact 2340
ggccctggag tacgcactgt gcctcacagg cgacggcttg gcccacctgc aggccaccga 2400
cccccagcag ctgctccgcc tcatccccca tgtgcaggtg ttcgcccgtg tggctcccaa 2460
gcagaaggag tttgtcatca ccagcctgaa ggagctgggc tacgtgaccc tcatgtgtgg 2520
ggatggcacc aacgacgtgg gcgccctgaa gcatgctgac gtgggtgtgg cgctcttggc 2580
caatgcccct gagcgggttg tcgagcggcg acggcggccc cgggacagcc caaccctgag 2640
caacagtggc atcagagcca cctccaggac agccaagcag cggtcggggc tccctccctc 2700
cgaggagcag ccaacctccc agagggaccg cctgagccag gtgctgcgag acctcgagga 2760
cgagagtacg cccattgtga aactggggga tgccagcatc gcagcaccct tcacctccaa 2820
gctctcatcc atccagtgca tctgccacgt gatcaagcag ggccgctgca cgctggtgac 2880
cacgctacag atgttcaaga tcctggcgct caatgccctc atcctggcct acagccagag 2940
cgtcctctac ctggagggag tcaagttcag tgacttccag gccaccctac aggggctgct 3000
gctggccggc tgcttcctct tcatctcccg ttccaagccc ctcaagaccc tctcccgaga 3060
acggcccctg cccaacatct tcaacctgta caccatcctc accgtcatgc tccagttctt 3120
tgtgcacttc ctgagccttg tctacctgta ccgtgaggcc caggcccgga gccccgagaa 3180
gcaggagcag ttcgtggact tgtacaagga gtttgagcca agcctggtca acagcaccgt 3240
ctacatcatg gccatggcca tgcagatggc caccttcgcc atcaattaca aggcccgcc 3300
cttcatggag agcctgcccg agaacaagcc cctggtgtgg agtctggcag tttcactcct 3360
ggccatcatt ggcctgctcc tcggctcctc gcccgacttc aacagccagt ttggcctcgt 3420
ggacatccct gtggagttca agctggtcat tgcccaggtc ctgctcctgg acttctgcct 3480
ggcgctcctg gccgaccgcg tcctgcagtt cttcctgggg accccgaagc tgaaagtgcc 3540
ttcctgagat ggcagtgctg gtacccactg cccaccctgg ctgccgctgg gcgggaaccc 3600
caacagggcc ccgggaggga accctgcccc caacccccca cagcaaggct gtacagtctc 3660
gcccttggaa gactgagctg ggacccccac agccatccgc tggcttggcc agcagaacca 3720
gccccaagcc agcacctttg gtaaataaag cagcatctga gattttaaaa 3770

<210> 108
<211> 1978
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510680CB1

<400> 108
cctcctacct gcggctctgg gccctcagcc tggctcatgc acaactgtct gaagtgctct 60
ggactatggt gatgaacagc ggccttcaga cgcgaggctg gggaggaatc gtcggggttt 120
ttattatttt tgccgtattt gctgtcctga cagtagccat ccttctgatc atggagggcc 180
tctctgcttt cctgcacgcc ctgcgactgc actgggaagc tgtttgggga acttgagcta 240
tttagaagat ggcaaccaag ccaacagagc ctgtcacgat cctcagcctt cggaaattga 300
gcctggggac cgcagagcca caggttaaag agccaaagac gttcaccgtg gaagatgcag 360
tggagactat cggcttcggg cgtttccaca ttgccctctt tctgatcatg ggcagtactg 420
gggtggttga ggccatggag atcatgttga tagctgttgt gtctcctgtc atccgctgtg 480
aatggcaact ggagaattgg caggtggcat tagtaaccac gatggtgttt tttggctaca 540
tggttttcag tatcctcttt ggcctcctgg ctgacagata tggccgctgg aagattctgc 600
tcatctcgtt cctgtgggga gcctatttct ccttgctgac ctcgtttgct ccttcgtaca 660
tctggtttgt cttcctgcgg acgatggtgg gctgtggtgt gtccggccac tcgcaagggt 720
taatcataaa gactgaattt ttgcccacga aataccgagg ctatatgtta cccttgtctc 780
aggtgttctg gcttgcgggc tccctgctca tcattggctt ggcctctgtg atcatcccca 840
ccatcgggtg gcgctggctc attcgcgtcg cctccatccc gggcatcatc ctcatcgtgg 900
ccttcaagtt tattcctgaa tctgcccggt tcaatgtctc cactgggaac actcgggctg 960
ccctggccac tctggagcgc gttgccaaga tgaaccgctc ggtcatgccg gaggggaagc 1020
tggtggagcc cgtcctggaa aaaagaggaa gatttgcaga cctattggat gctaaatatt 1080

tacggaccac attacagatc tgggtcatat ggcttggaat ctcttttgcc tactatgggg 1140
ttatcctggc cagtgctgag ctgctggagc gggacttggt ctgtggttca aagtcagact 1200
ctgcggtggt ggtgactggg ggggactcag gggagagcca gagcccctgc tactgccaca 1260
tgtttgcacc ctctgactat cggaccatga tcatcagcac catcggtgaa attgctttga 1320
atcctttaaa tatactgggc atcaatttcc tgggaagacg gctgagcctt tctattacca 1380
tgggatgcac ggctttattc ttccttctcc tcaacatttg cacttcaagt gccggcctga 1440
ttggcttcct cttcatgctg agggctctgg tagctgcaaa cttcaacacc gtctacattt 1500
acacagctga ggttcttatg agtgcatcaa tactgggggc cctgtgtctc ttctcatctg 1560
tctgtgttgt atgcgccatt tctgcattca ctctccccat cgaaaccaaa ggacgggccc 1620
tccagcaaat taaatgaaga cctgcaaagc tatgtctacc agatgagaaa aatgaattct 1680
atcttcagaa ctgcggtgca ttttttttaaa acttggtttt acttctgtat gctactcggt 1740
aattagtaaa gtgatttttt ttttaaaggc atatatggga atggggtagg taactgtata 1800
ttgatctctt ccttgaggaa caatatataa agtactttta taaaatataa tttaagcttt 1860
caaaggggtg tgagagggag atggtggggg ggaagatggc ttttcttcat tgaaatcaag 1920
tctgtaaacc tttatatgaa taaatactaa attttaaact tacaaaaaaa aaaaaagg 1978

<210> 109
<211> 1622
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505145CB1

<400> 109
tgcctggtgc cctgttccgc gtctgtgcga ccgtccgtcc cgagcgcgca ccgcggcagg 60
ccaggtcaga gcagcccacc atgggatggg gagggggtgg aggctgcacc ccgcgcccac 120
ccatccacca gcagccgccg gagcgccgcg tggtcaccgt tgtctttctc ggcctcctgc 180
tggacctcct ggccttcacg ctgctgctgc ccctgctgcc cgggctgttg gagagccacg 240
gccgtgccca cgacccccctc tatggctcct ggcagggcgg ggtggactgg tttgccaccg 300
ccatcgggat gccagtggag aagaggtaca acagtgtcct gttcggaggt ctcattggct 360
cggcattctc tgtcctgcag tttctgtgtg cgccactcac tggggccacc tctgactgct 420
tggggaggcg cccggtgatg ctgctgtgcc tgatgggtgt ggccacctca tatgcagtct 480
gggccacctc tcggagcttt gcggccttcc tggcctccag gctgattggg ggcatcagca 540
aagggaacgt cagcctctcc acggccatcg ttgctgacct gggctcgcct ctggcccgca 600
gtcaaggcat ggcggtcatt ggggtggcct tctcactggg cttcaccctg ggccctatgc 660
tcggagcctc cctgcccctg gaaatggcac cctggtttgc cctgctcttc gcagcctccg 720
acctgctgtt catcttctgc ttcctgccag agacgctgcc cctggagaaa cgggcgccct 780
ctatcgccct ggggttccgt gatgcggctg atctgctcag ccccctggcc ctgctgcgct 840
tctcggctgt cgctcgtggc caggacccac cctctggaga caggctcagc agcctgcgcc 900
gcctgggcct agtctacttc ctctacctct tcctgttctc gggcctggag tacacgctga 960
gcttcctcac acaccagcgc ttccagttca gtaggccctc ctgctgctgg tgcccgcctt 1020
cctcctcatc ggctggggac gttctctgcc cgtgctgggc ctggggctgc tgctctactc 1080
ctttgccgcc gccgttgtgg tgccctgcct gtcctccgtg gtcgctggct atggctcacc 1140
agggcagaag ggcacggtca tgggtacact gcgcagccta ggtgctctgg ccagggccgc 1200
ggggcccctg gtggccgctt cagtgtactg gctggccggg gcccaggcct gcttcaccac 1260
gtggtccggg ctcttttgc tccccttctt cctcctgcag aagctgagtt acccggcaca 1320
gacgctcaag gctgagtagc tgagccactg tgcccaggct gtgggcacca ggcagagtgg 1380
gagcctaggt caggcccctg cccactgcct gaccccccacc ccccgccagt ccagggagac 1440
cctgtgggtg gggccggcc cctaagcagg aagctcaggc agctcctcca gacttactta 1500
ctccttcagt gactccgagc tgcagcactc caaggctgtc agggcttctg tttgtttttt 1560
aaactatgca ccaggtttct gatgatgaaa taaagcacct gtttgtttta taaaaaaaaa 1620
aa 1622

<210> 110

<211> 1982
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505162CB1

<400> 110
ggctcccgcc tctgttcagg acactgggtc cccttggagc ctccccaggc ttaatgattg 60
tccagaaggc ggctataaag ggagcctggg aggctgggtg gaggagggag cagaaaaaac 120
ccaactcagc agatctggga actgtgagag cggcaagcag gaactgtggt cagaggctgt 180
gcgtcttggc tggtagggcc tgctcttttc taccatggca gcccagggct atggctatta 240
tcgcactgtg atcttctcag ccatgtttgg gggctacagc ctgtattact tcaatcgcaa 300
gaccttctcc tttgtcatgc catcattggt ggaagagatc cctttggaca aggatgattt 360
ggggttcatc accagcagcc agtcggcagc ttatgctatc agcaagtttg tcagtggggt 420
gctgtctgac cagatgagtg ctcgctggct cttctcttct gggctgctcc tggttggcct 480
ggtcaacata ttctttgcct ggagctccac agtacctgtc tttgctgccc tctggttcct 540
taatggcctg gcccaggggc tgggctggcc cccatgtggg aaggtcctgc ggaagtggtt 600
tgagccatct cagtttggca cttggtgggc catcctgtca accagcatga acctggctgg 660
agggctgggc cctatcctgg caaccatcct tgcccagagc tacagctggc gcagcacgct 720
ggccctatct ggggcactgt gtgtggttgt ctccttcctc tgtctcctgc tcatccacaa 780
tgaacctgct gatgttggac tccgcaacct ggaccccatg ccctctgagg gcaagaaggg 840
ctccttgaag gaggagagca ccctgcagga gctgctgctg tccccttacc tgtgggtgct 900
ctccactggt taccttgtgg tgtttggagt aaagacctgc tgtactgact ggggccagtt 960
cttccttatc caggagaaag gacagtcagc ccttgtaggc gggactgtcc aactacggga 1020
accctcgcca tggcctgttg ctgttcatga tggctggcat gacagtgtcc atgtacctct 1080
tccgggtaac agtgaccagt gactccccca agctctggat cctggtattg ggagctgtat 1140
ttggtttctc ctcgtatggc cccattgccc tgtttggagt catagccaac gagagtgccc 1200
ctcccaactt gtgtggcacc tcccacgcca ttgtgggact catggccaat gtgggcggct 1260
ttctggctgg gctgcccttc agcaccattg ccaagcacta cagttggagc acagccttct 1320
gggtggctga agtgatttgt gcggccagca cggctgcctt cttcctccta cgaaacatcc 1380
gcaccaagat gggccgagtg tccaagaagg ctgagtgaag agagtccagg ttccggagca 1440
ccatcccacg gtggccttcc ccctgcacgc tctgcgggga gaaaaggagg ggcctgcctg 1500
gctagccctg aacctttcac tttccatttc tgcgcctttt ctgtcacccg ggtggcgctg 1560
gaagttatca gtggctagtg aggtcccagc tccctgatcc tatgctctat ttaaaagata 1620
acctttggcc ttagactccg ttagctccta tttcctgcct tcagacaaac aggaaacttc 1680
tgcagtcagg aaggctcctg taccсttctt cttttcctag gccctgtcct gcccgcatcc 1740
taccccatcc ccacctgaag tgaggctatc cctgcagctg cagggcacta atgacccttg 1800
acttctgctg ggtcctaagt cctctcagca gtgggtgact gctgttgcca atacctcaga 1860
ctccagggaa agagaggagg ccatcattct cactgtacca ctaggcgcag ttggatatag 1920
gtgggaagaa aaggtgactt gttatagaag attaaaacta gatttgatac tgaaaaaaaa 1980
aa 1982

<210> 111
<211> 2231
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505469CB1

<400> 111
ggcttatgaa agtgtgatgt tcgcgtattt cttgccaggc agtggcgtga tcttggctca 60
ctgcaacctc cgactccctg gttcaagcga ttctcctgcc tcagcctcct gagtggggat 120

tacaggccac agcaaacaca ggtgtgcagg aaccgtttgt catggaagcc agggagcctg 180
ggaggcccac acccacctac catcttgtcc ctaacaccag ccagtcccag gtggaagaag 240
atgtcagctc gccacctcaa aggtcctccg aaactatgca gctgaagaag gagatctccc 300
tgctgaatgg ggtcagcctg gtggtgggca acatgatcgg ctcagggatc tttgtctcac 360
ccaagggtgt gctggtacac actgcctcct atgggatgtc actgattgtg tgggccattg 420
gtgggctctt ctctgttgtg ggtgcccttt gttatgcaga gctggggacc accatcacca 480
agtcgggagc cagctacgct tatattctag aggcctttgg gggcttcatt gccttcatcc 540
gcctgtgggt ctcactgcta gttgttgagc ccaccggtca ggccatcatc gccatcacct 600
ttgccaacta catcatccag ccgtccttcc ccagctgtga tcccccatac ctggcctgcc 660
gtctcctggc tgctgcttgc atatgtctgc tgacatttgt gaactgtgcc tatgtcaagt 720
ggggcacacg tgtgcaggac acgttcactt acgccaaggt cgtagcgctc attgccatca 780
ttgtcatggg ccttgttaaa ctgtgccagg aaatttgccc ttggccattg ggatttctat 840
gccaattgtg acgctcatct acatcctgac caatgtggcc tattacacag tgctgaacat 900
ttcagatgtc cttagcagtg atgctgtggc tgtgacattt gctgaccaga cgtttggcat 960
gttcagctgg accatcccca ttgctgttgc cctgtcctgc tttgggggcc tcaatgcatc 1020
catctttgct tcatcaaggt tgttcttcgt gggctcccgg gagggccacc taccggacct 1080
tctgtccatg atccacattg agcgttttac acctatccct gctttactgt tcaattgcac 1140
catggcactc atctacctca tcgtggagga tgttttccag cttatcaact acttcagctt 1200
cagctactgg ttcttcgtgg gcctgtctgt tgttggacag ctctacctcc gctggaagga 1260
gcccaagcgg ccccggcctc tcaagctgag cgtgtttttc cccatcgtgt tctgcatatg 1320
ctccgtgttt ctggtgatag tgcccctctt cactgacacc attaattccc tcattggcat 1380
cgggattgcc ctttctggag tccctttcta cttcatgggt gtttacctgc cagagtcccg 1440
gaggccattg tttattcgga atgtcctggc tgctatcacc agaggcaccc agcagctttg 1500
cttttgtgtc ctgactgagc ttgatgtagc cgaagaaaaa aaggatgaga ggaaaactga 1560
ctagaggtca gaggtggctt tctgaggcct ggaaggcagg ccaaccagca aaatcctgat 1620
aacaagactc tgtgggccca actctcctga attaaaggag ccttttgacc caatcatata 1680
gtggggctca gggccagtgc tcactcttat tggctttgtt tacacttgga gccaccttgg 1740
tgtgggtcac cgggacagtg tactcctctc ctgccagcct ccccttcccc gaggtgtggt 1800
ggctgcagtc tcaggaagag cttggtactt gtggggactt ctgttttctc cctgtggaga 1860
tcagtgaaga ctgggaggaa agctgcttca acctgagtcc ggctcttcag caggctgcac 1920
aagtggaagc aactaattct ggtgctcagg ctgggctctc cacccaagtt aggcctgctc 1980
tggcctaatg gatcttactg tatgagcagg acggctgcat tggattgtac aactgttttg 2040
tgatgccccc agacactgtc atcctaggcc gagaagaacc tgctagcttg acatacccca 2100
tgggcttatc cttaggtttt ggaattggtc aacagtgagg cagtctccct tcctgaccat 2160
tcttctccac ccagtcacag ataagggaat aaccttggcc atatatttgc tcaataaaga 2220
ttgaaggaag c 2231

<210> 112
<211> 5170
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505475CB1

<400> 112
agctgctgga gtaggcaccc atttaaagaa aaaatgaaga agcagcaata aagaagttgt 60
aatcgttacc tagacaaaca gagaactggt tttgacagtg tttctagagt gctttttatt 120
attttcctga cagttgtgtt cccaccatgat tactttctcc ttcagcgaat aggctaaatg 180
aatatgaaac agaaaagcgt gtatcagcaa accaaagcac ttctgtgcaa gaattttctt 240
aagaaatgga ggatgaaaag agagagctta ttggaatggg gcctctcaat acttctagga 300
ctgtgtattg ctctgttttc cagttccatg agaaatgtcc agtttcctgg aatggctcct 360
cagaatctgg gaagggtaga taaatttaat agctcttctt taatggttgt gtatacacca 420
atatctaatt taacccagca gataatgaat aaaacagcac ttgctcctct tttgaaagga 480
acaagtgtca ttggggcacc aaataaaaca cacatggacg aaatacttct ggaaaattta 540

ccatatgcta tgggaatcat ctttaatgaa actttctctt ataagttaat atttttccag 600
ggatataaca gtccactttg gaaagaagat ttctcagctc attgctggga tggatatggt 660
gagttttcat gtacattgac caaatactgg aatagaggat ttgtggcttt acaaacagct 720
attaatactg ccattataga aatcacaacc aatcaccctg tgatggagga gttgatgtca 780
gttactgcta taactatgaa gacattacct ttcataacta aaaatcttct tcacaatgag 840
atgtttattt tattcttctt gcttcatttc tccccacttg tatattttat atcactcaat 900
gtaacaaaag agagaaaaaa gtctaagaat ttgatgaaaa tgatgggtct ccaagattca 960
gcattctggc tctcctgggg tctaatctat gctggcttca tctttattat ttccatattc 1020
attacaatta tcataacatt cacccaaatt atagtcatga ctggcttcat ggtcatattt 1080
ataccctttt ttttatatgg cttatctttg gtagctttgg tgttcctgct gagtgtgctg 1140
ttaaagaaag ctgtcctcac caatttggtt gtgtttctcc ttaccctctt ttggggatgt 1200
ctgggattca ctgtatttta tgaacaactt ccttcatctc tggagtggat tttgaatatt 1260
tgtagccctt ttgcctttac tactggaatg attcagatta tcaaactgga ttataacttg 1320
aatggtgtaa tttttcctga cccttcagga gactcatata caatgatagc aactttttct 1380
atgttgcttt tggatggtct catctacttg ctattggcat tatactttga caaaatttta 1440
ccctatggag atgagcgcca ttattctcct ttatttttct tgaattcatc atcttgtttc 1500
caacaccaaa ggactaatgc taaggttatt gagaaagaaa tcgatgctga gcatccctct 1560
gatgattatt ttgaaccagt agctcctgaa ttccaaggaa aagaagccat cagaatcaga 1620
aatgttaaga aggaatataa aggaaaatct ggaaaagtgg aagcattgaa aggcttgctc 1680
tttgacatat atgaaggtca aatcacggca atcctgggtc acagtggagc tggcaaatct 1740
tcactgctaa atattcttaa tggattgtct gttccaacag aaggatcagt taccatctat 1800
aataaaaatc tctctgaaat gcaagacttg gaggaaatca gaaagataac tggcgtctgt 1860
cctcaattca atgttcaatt tgacatactc accgtgaagg aaaacctcag cctgtttgct 1920
aaaataaaag ggattcatct aaaggaagtg gaacaagagg tacaacgaat attattggaa 1980
ttggacatgc aaaacattca agataacctt gctaaacatt taagtgaagg acagaaaaga 2040
aagctgactt ttgggattac cattttagga gatcctcaaa tagaaaagtg atcatgtcca 2100
atgggagact gaagtgtgca ggttcttcta tgtttttgaa aagaaggtgg ggtcttggat 2160
atcacctaag tttacatagg aatgaaatat gtaacccaga acaaataaca tccttcatta 2220
ctcatcacat ccccgatgct aaattaaaaa cagaaaacaa agaaaagctt gtatatactt 2280
tgccactgga aaggacaaat acatttccag atcttttcag tgatctggat aagtgttctg 2340
accagggagt gacaggttat gacatttcca tgtcaactct aaatgaagtc tttatgaaac 2400
tggaaggaca gtcaactatc gaacaagatt tcgaacaagt ggagatgata agagactcag 2460
aaagcctcaa tgaaatggag ctggctcact cttccttctc tgaaatgcag acagctgtga 2520
gtgacatggg cctctggaga atgcaagtct ttgccatggc acggctccgt ttcttaaagt 2580
taaaacgtca aactaaagtg ttattgaccc tattattggt atttggaatc gcaatattcc 2640
ctttgattgt tgaaaatata atatatgcta tgttaaatga aaagatcgat tgggaattta 2700
aaaacgaatt gtattttctc tctcctggac aacttcccca ggaaccccgt accagcctgt 2760
tgatcatcaa taacacagaa tcaaatattg aagattttat aaaatcactg aagcatcaaa 2820
atatactttt ggaagtagat gactttgaaa acagaaatgg tactgatggc ctctcataca 2880
atggagctat catagtttct ggtaaacaaa aggattatag attttcagtt gtgtgtaata 2940
ccaagagatt gcactgtttt ccaattctta tgaatattat cagcaatggg ctacttcaaa 3000
tgtttaatca cacacaacat attcgaattg agtcaagccc atttcctctt agccacatag 3060
gactctggac tgggttgccg gatggttcct ttttcttatt tttggttcta tgtagcattt 3120
ctccttatat caccatgggc agcatcagtg attacaagaa aaatgctaag tcccagctat 3180
ggatttcagg cctctacact tctgcttact ggtgtgggca ggcactagtg gacgtcagct 3240
tcttcatttt aattctcctt ttaatgtatt taatttcta catagaaaac atgcagtacc 3300
ttcttattac aagccaaatt gtgtttgctt tggttatagt tactcctggt tatgcagctt 3360
ctcttgtctt cttcatatat atgatatcat ttatttttcg caaaaggaga aaaaacagtg 3420
gcctttggtc attttacttc ttttttgcct ccaccatcat gttttccatc actttaatca 3480
atcattttga cctaagtata ttgattacca ccatggtatt ggttccttca tataccttgc 3540
ttggatttaa aactttttg gaagtgagag accaggagca ctacagagaa tttccagagg 3600
caaatttga attgagtgcc actgattttc tagtctgctt catacccctac tttcagactt 3660
tgctattcgt ttttgttcta agatgcatgg aactaaaatg tggaaagaaa agaatgcgaa 3720
aagatcctgt tttcagaatt tccccccaaa gtagagatgc taagccaaat ccagaagaac 3780
ccatagatga agatgaagat attcaaacag aaagaataag aacagccact gctctgacca 3840
cttcaatctt agatgagaaa cctgttataa ttgccagctg tctacacaaa gaatatgcag 3900

gccagaagaa aagttgcttt tcaaagagga agaagaaaat agcagcaaga aatatctctt 3960
tctgtgttca agaaggtgaa attttgggat tgctaggacc cagtggtgct ggaaaaagtt 4020
catctattag aatgatatct gggatcacaa agccaactgc tggagaggtg gaactgaaag 4080
gctgcagttc agttttgggc cacctggggt actgccctca agagaacgtg ctgtggccca 4140
tgctgacgtt gagggaacac ctggaggtgt atgctgccgt caaggggctc aggaaagcgg 4200
acgcgaggct cgccatcgca agattagtga gtgctttcaa actgcatgag cagctgaatg 4260
ttcctgtgca gaaattaaca gcaggaatca cgagaaagtt gtgttttgtg ctgagcctcc 4320
tgggaaactc acctgtcttg ctcctggatg aaccatctac gggcatagac cccacagggc 4380
agcagcaaat gtggcaggca atccaggcag tcgttaaaaa cacagagaga ggtgtcctcc 4440
tgaccaccca taacctggct gaggcggaag ccttgtgtga ccgtgtggcc atcatggtgt 4500
ctggaaggct tagatgcatt ggctccatcc aacacctgaa aaacaaactt ggcaaggatt 4560
acattctaga gctaaaagtg aaggaaacgt ctcaagtgac tttggtccac actgagattc 4620
tgaagctttt cccacaggct gcagggcagg aaaggtattc ctctttgtta acctataagc 4680
tgcccgtggc agacgtttac cctctatcac agacctttca caaattagaa gcagtgaagc 4740
ataactttaa cctggaagaa tacagccttt ctcagtgcac actggagaag gtattcttag 4800
agctttctaa agaacaggaa gtaggaaatt ttgatgaaga aattgataca acaatgagat 4860
ggaaactcct ccctcattca gatgaacctt aaaacctcaa acctagtaat tttttgttga 4920
tctcctataa acttatgttt tatgtaataa ttaatagtat gtttaatttt aaagatcatt 4980
taaaattaac atcaggtata ttttgtaaat ttagttaaca aatacataaa ttttaaaatt 5040
attcttcctc tcaaacatag gggtgatagc aaacctgtga taaaggcaat acaaaatatt 5100
agtaaagtca cccaaagatt caggcactgg aattatggac taaaacggta tatctagttt 5160
tagcctcttc 5170

<210> 113
<211> 1876
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7505568CB1

<400> 113
gccttggcag agtctggggt ccctggactg agccatcagc tgggtcactg agacccatgg 60
caaggaaaca aaataggaat tccaaggaac tgggcctagt tcccctcaca gatgacacca 120
gccacgccgg gcctccaggg ccagggaggg cactgctgga gtgtgaccac ctgaggagtg 180
gggtgccagg tggaaggaga agaaaggact ggtcctgctc gctcctcgtg gcctccctcg 240
cgggcgcctt cggctcctcc ttcctctacg gctacaacct gtcggtggtg aatgcccccca 300
ccccggaagc acactttgct ggccaataat gggtttgcaa tttctgctgc attgctgatg 360
gcctgctcgc tccaggcagg agcctttgaa atgctcatcg tgggacgctt catcatgggc 420
atagatggag gcgtcgccct cagtgtgctc cccatgtacc ttagtgagat ctcacccaag 480
gagatccgtg gctctctggg gcaggtgact gccatcttta tctgcattgg cgtgttcact 540
gggcagcttc tgggcctgcc cgagctgctg ggaaaggaga gtacctggcc atacctgttt 600
ggagtgattg tggtccctgc cgttgtccag ctgctgagcc ttccctttct cccggacagc 660
ccacgctacc tgctcttgga gaagcacaac gaggcaagag ctgtgaaagc cttccaaacg 720
ttcttgggta aagcagacat ttcccaagag gtagaggagg tcctggctga gagccgcgtg 780
cagaggagca tccgcctggt gtccgtgctg gagctgctga gagctcccta cgtccgctgg 840
caggtggtca ccgtgattgt caccatggcc tgctaccagc tctgtggcct caatgcaatt 900
tggttctata ccaacagcat ctttggaaaa gctgggatcc ctccggcaaa gatcccatac 960
gtcaccttga gtacaggggg catcgagact ttggctgccg tcttctctgg tttggtcatt 1020
gagcacctgg gacggagacc cctcctcatt ggtggctttg ggctcatggg cctcttcttt 1080
gggaccctca ccatcacgct gaccctgcag gaccacgccc ccctgggtcc cctacctgag 1140
tatcgtgggc attctggcca tcatcgcctc tttctgcagt gggcccaggt ggcatcccgt 1200
tcatcttgac tggtgagttc ttccagcaat ctcagcggcc ggctgccttc atcattgcag 1260
gcaccgtcaa ctggctctcc aactttgctg ttgggctcct cttcccattc attcagaaaa 1320
gtctggacac ctactgtttc ctagtctttg ctacaatttg tatcacaggt gctatctacc 1380

tgtattttgt gctgcctgag accaaaaaca gaacctatgc agaaatcagc caggcatttt 1440
ccaaaaggaa caaagcatac ccaccagaag agaaaatcga ctcagctgtc actgatgcag 1500
ctcaaaggaa ctaagacaaa gatcatggag accatcgggt gagtctcaag acttccccca 1560
gctctgcttg gctggtctcc tgctggtatt ttctgtctgt agagaggaac aagaacttcc 1620
attttatctt gcttacctgc acttatgaaa agtcaaactg agtcatgctg agagccaggg 1680
aacataggag tcagttcttc tgcagcagca ctcagccagt tgaaggcaat gtggagtgat 1740
ggaaggagag cagtgatgca gtgatgctgg caccaactcc tttactatgg catccattgt 1800
accagctgcc atacaccagg caacttctac actttatctc taatcatcct agaataagta 1860
ttagtttccc catctt 1876

<210> 114
<211> 1602
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7506953CB1

<400> 114
ctctctctct ctctctctct ctctctctct ctctctctct ctcccaagtt tcctatctcg 60
tcaagatcag ggcaaaagaa gaaaacaccg aattctgctt gccgtttcag agcggcggtg 120
atgaagacaa aattgaacat ctacaacatg cagttcctgc tttttgtttt cttggtgtgg 180
gaccctgcca ggttggtgct ggctaacatc caagaagatg aggctaaaaa taacattacc 240
atctttacga gaattcttga cagacttctg gatggttacg ataatcggct tagaccagga 300
ctgggagatg catatacaac ttcagaggtc acttatattt ggacttacaa tgcatctgat 360
tcagtacagg ttgctcctga tggctctagg ttaaatcaat atgacctgct gggccaatca 420
atcggaaagg agacaattaa atccagtaca ggtgaatata ctgtaatgac agctcatttc 480
cacctgaaaa gaaaaattgg gtattttgtg attcaaacct atctgccttg catcatgact 540
gtcattctct cccaagtttc attctggctt aacagagaat ctgtgcctgc aagaactgtg 600
tttggagtaa caactgtcct aacaatgaca actctaagca tcagtgctcg gaattctctc 660
cccaaagtgg cttatgcaac tgccatggac tggtttattg ctgtttgtta tgcatttgtg 720
ttctctgccc taattgaatt tgcaactgtt aattacttca ccaaaagagg atgggcttgg 780
gatgggaaga gtgtagtaaa tgacaagaaa aagaaaaagg cttccgttat gatacagaac 840
aacgcttatg cagtggctgt tgccaattat gccccgaatc tttcaaaaga tccagttctc 900
tccaccatct ccaagagtgc aaccacgcca gaacccaaca agaagccaga aaacaagcca 960
gctgaagcaa agaaaacttt caacagtgtt agcaaaattg acagaatgtc cagaatagtt 1020
tttccagttt tgtttggtac ctttaattta gtttactggg ctacatattt aaacagagaa 1080
cctgtattag ggtcagtcc ttgaattgag acccatgtta tctttgggat gtatagcaac 1140
attaaatttg gtttgttttg ctatgtacag tctgactaat aactgctaat ttgtgatcca 1200
acatgtacag tatgtatata gtgacatagc ttaccagtag acctttaatg gagacatgca 1260
tttgctaact catggaactg cagacagaaa gcactccatg cgaaaacagc cattgccttt 1320
tttaaagatt taccctagga cctgatttaa agtgaatttc aaatgacctg attaatttcc 1380
tattcttcca aatgagatga aaatggggat cctgtacaac cctttgtgga cacttttggt 1440
ttagctctta agtaggggta ttttctactg ttgcctaata tgatggaagt aacattgtca 1500
ttctagatga atctttgaag taccaacatt gttctggaat cagctcggtc cagatgctca 1560
aggtccctgt aatgtattgg aagctggtac cctaagaaaa ca 1602

<210> 115
<211> 2173
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510176CB1

<400> 115
cactattgaa gccacctgct caggacaatg aaattcttca gttacattct ggtttatcgc 60
cgatttctct tcgtggtttt cactgtgttg gttttactac ctctgcccat cgtcctccac 120
accaagctta ttctgacttt tccaagatga agatgaagag caggccagtg gcatgggcca 180
gatgcagtca gatgactgac tgaccaggag ggcctccgtc cagtgaccag tctcagaatt 240
taaaggaccc ttcccctta cttatgcttg tactcagccc aggaagcaga atgtgcctac 300
acactctttg tggtcgccac attttggctc acagaagcat tgcctctgtc ggtaacagct 360
ttgctaccta gtttaatgtt acccatgttt gggatcatgc cttctaagaa ggtggcatct 420
gcttatttca aggattttca cttactgcta attggagtta tctgtttagc aacatccata 480
gaaaaatgga atttgcacaa gagaattgct ctgaaaatgg tgatgatggt tggtgtaaat 540
cctgcatggc tgacgctggg gttcatgagc agcactgcct ttttgtctat gtggctcagc 600
aacacctcga cggctgccat ggtgatgccc attgcggagg ctgtagtgca gcagatcatc 660
aatgcagaag cagaggtcga ggccactcag atgacttact tcaacggatc aaccaaccac 720
ggactagaaa ttgatgaaag tgttaatgga catgaaataa atgagaggaa agagaaaaca 780
aaaccagttc caggatacaa taatgataca gggaaaattt caagcaaggt ggagttggaa 840
aagaactcag gcatgagaac caaatatcga acaaagaagg gccacgtgac acgtaaactt 900
acgtgtttgt gcattgccta ctcttctacc attggtggac tgacaacaat cactggtacc 960
tccaccaact tgatctttgc agagtatttc aatacacgct atcctgactg tcgttgcctc 1020
aactttggat catggtttac gttttccttc ccagctgccc ttatcattct actcttatcc 1080
tggatctggc ttcagtggct tttcctagga ttcaatttta aggagatgtt caaatgtggc 1140
aaaaccaaaa cagtccaaca aaaagcttgt gctgaggtga ttaagcaaga ataccaaaag 1200
cttgggccaa taaggtatca agaaattgtg accttggtcc tcttcattat aatggctctg 1260
ctatggttta gtcgagaccc cggatttgtt cctggttggt ctgcactttt ttcagagtac 1320
cctggttttg ctacagattc aactgttgct ttacttatag ggctgctatt ctttcttatc 1380
ccagctaaga cactgactaa aactacacct acaggagaaa ttgttgcttt tgattactct 1440
ccactgatta cttggaaaga attccagtca ttcatgccct gggatatagc cattcttgtt 1500
ggtggagggt ttgccctggc agatggttgt gaggagtctg gattatctaa gtggatagga 1560
aataaattat ctcctctggg ttcattacca gcatggctaa taattctgat atcttctttg 1620
atggtgacat ctttaactga ggtagccagc aatccagcta ccattacact ctttctccca 1680
atattatctc cattggccga agccattcat gtgaaccctc tttatattct gataccttct 1740
actctgtgta cttcatttgc attcctccta ccagtagcaa atccacccaa tgctattgtc 1800
ttttcatatg gtcatctgaa agtcattgac atggttaaag ctggacttgg tgtcaacatt 1860
gttggtgttg ctgtggttat gcttggcata tgtacttgga ttgtacccat gtttgacctc 1920
tacacttacc cttcgtgggc tcctgctatg agtaatgaga ccatgccata ataagcacaa 1980
aatttctgac tatcttgcgg taatttctgg aagacattaa tgattgactg taaaatgtgg 2040
ctctaaataa ctaatgacac acatttaaat cagttatggt gtagctgctg caattcccgt 2100
gaatacccga aacctgctgt tataactcag agtccatatt tgttattgca gtgcaactaa 2160
agagcatcta tgt 2173

<210> 116
<211> 1826
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510541CB1

<400> 116
cccacgcgtc cgcccacgcg tccgcccacg cgtccgcagc cgagttcagc cgggagcaga 60
gcgaaccgca ccggcccgag cggagcgccg cacgttccca accgcgaggc cagacatctg 120
actgttggtg tgagaccagt gctcctggtg gtgtgccctg agccatggag gcgcctttgc 180
agacagagat ggtggagctg gtgcccaatg gcaaacactc agaggggctg ctcccggtca 240
tcaccccat ggcaggcaac cagagggtcg aggaccctgc acggagctgt atggagggca 300
agagcttcct acagaaaagt cccagcaagg agccacactt cactgacttc gaggggaaga 360
catcattcgg gatgtcagtg ttcaacctca gcaatgccat catgggcagc ggcatcctgg 420

gactcgccta tgccatggcc aatacgggca ttatcctttt cctgcatccg tgcctatgag 480
cagctgggct accgtgcctt tgggacccca ggaaagctgg cagcagcctt ggccatcacg 540
ctccagaaca tcggagccat gtccagctac ctgtacatca tcaagtctga gctgccactt 600
gtcatacaga ccttcctgaa cctggaggag aaaacctcgg actggtacat gaacgggaac 660
tacctggtaa tccttgtctc tgtcaccatc attctgcccc tggcactgat gcggcagctt 720
ggctacctgg gctactccag cggcttctct cttagctgca tggtgttctt cctaattgca 780
gtcatctaca aaaagttcca cgtgccctgc ccactgcccc ccaacttcaa caacaccaca 840
ggcaacttca gccacgtgga gatcgtgaag gagaaggtgc agctgcaggt cgagcctgag 900
gcttcagcct tctgcactcc cagctacttc acgctcaact cacagacagc atacaccatc 960
cccatcatgg ccttcgcctt cgtctgccac cccgaggtgc tgcccatcta tactgagctc 1020
aaggacccct ccaagaagaa gatgcagcac atctccaacc tgtccatcgc tgtcatgtac 1080
atcatgtact tcctggctgc cctcttcggc tacctcacct tctacaacgg ggtggagtcg 1140
gagctgctgc acacctacag caaggtggac ccgtttgacg tcctgatcct gtgtgtgcgc 1200
gtggccgtgc tgacagcagt cacgctcaca gtgcccatcg ttctgttccc ggtgcgccgc 1260
gccatccagc agatgctgtt tccaaaccag gagttcagct ggctgcggca tgtgcttatt 1320
gccgttggcc tgctcacttg tatcaacctg ctggtcatct ttgcccccaa catcctgggc 1380
atctttgggg tcatcggtgc cacatctgcc ccattcctca tcttcatctt ccctgccatc 1440
ttctacttcc gaatcatgcc cacggagaag gagcctgcaa gatccacccc caaaatcctg 1500
gccctgtgtt ttgctatgct tggcttcttg ctgatgacca tgagcttgag cttcatcatc 1560
attgactggg cctcagggac cagccggcat ggaggaaacc actagggtga ccctcatcct 1620
gttctgtcta ctcaccctag cagccctgcc cagactcttc agcccctgct cccatccagt 1680
ggccagtcgg gggaggagaa agacgcgatt aacactgtgg cattcagcca ggccccatgt 1740
cctctctgtg gaaggttttt gttcaagagc caggaccaag gcccttgggc cactaccctg 1800
ctaggctctg gagctgtaga ggcttc 1826

<210> 117
<211> 2052
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510923CB1

<400> 117
cccacgcgtc cgcccacgcg tccgcccacg cgtccgcagc cgagttcagc cgggagcaga 60
gcgaaccgca ccggcccgag cggagcgccg cacgttgcca accgcgaggc cagacatctg 120
actgttggtg tgagaccagt gctcctggtg gtgtgccctg agccatggag gcgcctttgc 180
agacagagat ggtggagctg gtgcccaatg gcaaacactc agaggggctg ctcccggtca 240
tcacccccat ggcaggcaac cagagggtcg aggaccctgc acggagctgt atggagggca 300
agagcttcct acagaaaagt cccagcaagg agccacactt cactgacttc gaggggaaga 360
catcattcgg gatgtcagtg ttcaacctca gcaatgccat catgggcagc ggcatcctgg 420
gactcgccta tgccatggcc aatacgggca ttatcctttt cctgttcctg ttgacagctg 480
tcgccttgct ctccagctac tccatccacc tgctactcaa gtcctcaggg gtcgtgggca 540
tccgtgccta tgagcagctg ggctaccgtg cctttgggac cccaggaaag ctggcagcag 600
ccctggccat cacgctccag aacatcggag ccatgtccag ctacctgtac atcatcaagt 660
ctgagctgcc acttgtcata cagaccttcc tgaacctgga ggagaaaacc tcggactggt 720
acatgaacgg gaactacctg gtaatccttg tctctgtcac catcattctg cccctggcac 780
tgatgcggca gcttggctac ctgggctact ccagcggctt ctctcttagc tgcatggtgt 840
tcttcctaat tgcagtcatc tacaaaaagt tccacgtgcc ctgcccactg ccccccaact 900
tcaacaacac cacaggcaac ttcagccacg tggagatcgt gaaggagaag gtgcagctgc 960
aggtcgagcc tgaggcttca gccttctgca ctcccagcta cttcacgctc aactcacagg 1020
ttctgacagg tcagggcaag gcgggggccc aatgagagtg gcagactgcc ttgacacagc 1080
cccgtgtttc cccagacagc atacaccatc cccatcatgg ccttcgcctt cgtctgccac 1140
cccgaggtgc tgcccatcta tactgagctc aaggacccct ccaagaagaa gatgcagcac 1200
atctccaacc tgtccatcgc tgtcatgtac atcatgtact tcctggctgc cctcttcggc 1260

tacctcacct tctacaacgg ggtggagtcg gagctgctgc acacctacag caaggtggac 1320
ccgtttgacg tcctgatcct gtgtgtgcgc gtggccgtgc tgacagcagt cacgctcaca 1380
gtgcccatcg ttctgttccc ggtgcgccgc gccatccagc agatgctgtt tccaaaccag 1440
gagttcagct ggctgcggca tgtgcttatt gccgttggcc tgctcacttg tatcaacctg 1500
ctggtcatct ttgcccccaa catcctgggc atctttgggg tcatcggtgc cacatctgcc 1560
ccattcctca tcttcatctt ccctgccatc ttctacttcc gaatcatgcc cacggagaag 1620
gagcctgcaa gatccacccc caaaatcctg gccctgtgtt ttgctatgct tggcttcttg 1680
ctgatgacca tgagcttgag cttcatcatc attgactggg cctcagggac cagccggcat 1740
ggaggaaacc actagggtga ccctcatcct gttctgtcta ctcaccctag cagccctgcc 1800
cagactcttc agcccctgct cccatccagt ggccagtcgg gggaggagaa agacgcgatt 1860
aacactgtgg cattcagcca ggccccatgt cctctctgtg gaaggttttt gttcaagagc 1920
caggaccaag gcccttgggc cactaccctg ctaggctctg gagctgtaga ggcttcctga 1980
aactgggagc agggtagggc tgtcgccttt agatcccgcc caagccccct cattccctcc 2040
tttgcacaga tg 2052

<210> 118
<211> 5056
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7510984CB1

<400> 118
cgcgcggagc cagcggagcc agctgagccc gagcccagcc cgcgcccgcg ccgccatgcc 60
cctggccttc tgcggcagcg agaaccactc ggccgcctac cgggtggacc agggggtcct 120
caacaacggc tgctttgtgg acgcgctcaa cgtggtgccg cacgtcttcc tactcttcat 180
caccttcccc atcctcttca ttggatgggg aagtcagagc tccaaggtgc acatccacca 240
cagcacatgg cttcatttcc ctgggcacaa cctgcggtgg atcctgacct tcatgctgct 300
cttcgtcctg gtgtgtgaga ttgcagaggg catcctgtct gatggggtga ccgaatccca 360
ccatctgcac ctgtacatgc cagccgggat ggcgttcatg gctgctgtcg cctccgtggt 420
ctactatcac aacatcgaga cttccaactt ccccaagctg ctaattgccc tgctggtgta 480
ttggaccctg gccttcatca ccaagaccat caagtttgtc aagttcttgg accacgccat 540
cggcttctcg cagctacgct tctgcctcac agggctgctg gtgatcctct atgggatgct 600
gctcctcgtg gaggtcaatg tcatcagggt gaggagatac atcttcttca agacaccgag 660
ggaggtgaag cctcccgagg acctgcaaga cctgggggta cgcttcctgc agcccttcgt 720
gaatctgctg tccaaaggca cctactggtg gatgaacgcc ttcatcaaga ctgcccacaa 780
gaagcccatc gacttgcgag ccatcgggaa gctgcccatc gccatgaggg ccctcaccaa 840
ctaccaacgg ctctgcgagg cctttgacgc ccaggtgcgg aaggacattc agggcactca 900
aggtgcccgg gccatctggc aggcactcag ccatgccttc gggaggcgcc tggtcctcag 960
cagcactttc cgcatcttgg ccgacctgct gggcttcgcc gggccactgt gcatctttgg 1020
gatcgtggac caccttggga aggagaacga cgtcttccag cccaagacac aatttctcgg 1080
ggtttacttt gtctcatccc aagagttcct tgccaatgcc tacgtcttag ctgtgcttct 1140
gttccttgcc ctcctactgc aaaggacatt tctgcaagca tcctactatg tggccattga 1200
aactggaatt aacttgagag gagcaataca gaccaagatt tacaataaaa ttatgcacct 1260
gtccacctcc aacctgtcca tgggagaaat gactgctgga cagatctgta atctggttgc 1320
catcgacacc aatcagctca tgtggttttt cttcttgtgc ccaaacctct gggctatgcc 1380
agtacagatc attgtgggtg tgattctcct ctactacata ctcggagtca gtgccttaat 1440
tggagcagct gtcatcattc tactggctcc tgtccagtac ttcgtggcca ccaagctgtc 1500
tcaggcccag cggagcacac tggagtattc caatgagcgg ctgaagcaga ccaacgagat 1560
gctccgcggc atcaagctgc tgaagctgta cgcctgggag aacatcttcc gcacgcgggt 1620
ggagacgacc cgcaggaagg agatgaccag cctcagggcc tttgccatct atacctccat 1680
ctccattttc atgaacacgg ccatccccat tgcagctgtc ctcataactt tcgtgggcca 1740
cgtcagcttc ttcaaagagg ccgacttctc gccctccgtg gcctttgcct ccctctccct 1800
cttccatatc ttggtcacac cgctgttcct gctgtccagt gtggtccgat ctaccgtcaa 1860

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 October 2003 (09.10.2003)

PCT

(10) International Publication Number
WO 2003/083085 A3

(51) International Patent Classification:
*A61K 38/00* (2006.01) *C07K 16/00* (2006.01)
*A61K 39/00* (2006.01) *C12N 1/20* (2006.01)
*C07H 21/04* (2006.01) *C12N 5/00* (2006.01)
*C07K 1/00* (2006.01) *C12N 15/09* (2006.01)
*C07K 14/00* (2006.01) *C12P 21/00* (2006.01)

(21) International Application Number: PCT/US2003/009797

(22) International Filing Date: 27 March 2003 (27.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/368,840 28 March 2002 (28.03.2002) US
60/375,637 26 April 2002 (26.04.2002) US

(71) Applicant *(for all designated States except US)*: **INCYTE CORPORATION** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MARQUIS, Joseph, P.** [US/US]; 4428 Lazy Lane, San Jose, CA 95135 (US). **LEE, Soo, Y.** [KR/US]; 40 Westdale Avenue, Daly City, CA 94015 (US). **EMERLING, Brooke, M.** [US/US]; 833 W. Buena Avenue #2108, Chicago, IL 60613 (US). **HAFALIA, April, J.A.** [US/US]; 15 Midvale Drive, Daly City, CA 94015 (US). **KHARE, Reena** [IN/US]; 12650 Orella Court, Saratoga, CA 95070 (US). **KABLE, Amy, E.** [US/US]; 2345 Polk Street #4, San Francisco, CA 94109 (US). **RICHARDSON, Thomas, W.** [US/US]; 616 Canyon Road #107, Redwood City, CA 94062 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue #5D, San Francisco, CA 94122 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square, #712, Union City, CA 94587 (US). **BECHA, Shanya, D.** [US/US]; 1 Saint Francis #5508, San Francisco, CA 94107 (US). **MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95014 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **TRAN, Uyen, K.** [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US). **ISON, Craig, H.** [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). **LINDQUIST, Erika, A.** [US/US]; 2394 Mariner Square Drive #C-121, Alameda, CA 94501 (US). **JIANG, Xin** [US/US]; 14371 Elva Avenue, Saratoga, CA 95070 (US). **JACKSON, Alan, A.** [US/US]; 1541 Elwood Drive, Los Gatos, CA 95032 (US). **WILSON, Amy, D.** [US/US]; 1056 Continentals Way, #27, Belmont, CA 94002 (US). **JIN, Pei** [US/US]; 320 Curtner Avenue #D, Palo Alto, CA 94306 (US). **CHANG, Hsin-Ru** [US/US]; 326 Treasure Island Drive, Belmont, CA 94002 (US).

(74) Agent: **FOLEY & LARDNER LLP**; Washington Harbour, 3000 K Street, N.W., Suite 500, Washington, D.C. 20007-5143 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report: 23 March 2006

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: TRANSPORTERS AND ION CHANNELS

(57) Abstract: Various embodiments of the invention provide human transporters and ion channels (TRICH) and polynucleotides which identify and encode TRICH. Embodiments of the invention also provide expression vectors, host cells, antibodies, agonists, and antagonists. Other embodiments provide methods for diagnosing, treating, or preventing disorders associated with aberrant expression of TRICH.

WO 2003/083085 A3

agctctagtg agcgtgcaaa agctaagcga gttcctgtcc agtgcagaga tccgtgagga 1920
gcagtgtgcc ccccatgagc ccacacctca gggcccagcc agcaagtacc aggcggtgcc 1980
cctcagggtt gtgaaccgca agcgtccagc ccgggaggat tgtcggggcc tcaccggccc 2040
actgcagagc ctggtcccca gtgcagatgg cgatgctgac aactgctgtg tccagatcat 2100
gggaggctac ttcacgtgga ccccagatgg aatccccaca ctgtccaaca tcaccattcg 2160
tatcccccga ggccagctga ctatgatcgt ggggcaggtg ggctgcggca agtcctcgct 2220
ccttctagcc gcactggggg agatgcagaa ggtctcaggg gctgtcttct ggagcagcag 2280
ccttcctgac agcgagatag gagaggaccc cagcccagag cgggagacag cgaccgactt 2340
ggatatcagg aagagaggcc ccgtggccta tgcttcgcag aaaccatggc tgctaaatgc 2400
cactgtggag gagaacatca tctttgagag tcccttcaac aaacaacggt acaagatggt 2460
cattgaagcc tgctctctgc agccagacat cgacatcctg ccccatggag accagaccca 2520
gattggggaa cggggcatca acctgtctgg tggtcaacgc cagcgaatca gtgtggcccg 2580
agccctctac cagcacgcca acgttgtctt cttggatgac cccttctcag ctctggatat 2640
ccatctgagt gaccacttaa tgcaggccgg catccttgag ctgctccggg acgacaagag 2700
gacagtggtc ttagtgaccc acaagctaca gtacctgccc catgcagact ggatcattgc 2760
catgaaggat ggcaccatcc agagggaggg taccctcaag gacttccaga ggtctgaatg 2820
ccagctcttt gagcactgga agaccctcat gaaccgacag gaccaagagc tggagaagga 2880
gactgtcaca gagagaaaag ccacagagcc accccagggc ctatctcgtg ccatgtcctc 2940
gagggatggc cttctgcagg atgaggaaga ggaggaagag gaggcagctg agagcgagga 3000
ggatgacaac ctgtcgtcca tgctgcacca gcgtgctgag atcccatggc gagcctgcgc 3060
caagtacctg tcctccgccg gcatcctgct cctgtcgttg ctggtcttct cacagctgct 3120
caagcacatg gtcctggtgg ccatcgacta ctggctggcc aagtggaccg acagcgccct 3180
gaccctgacc cctgcagcca ggaactgctc cctcagccag gagtgcaccc tcgaccagac 3240
tgtctatgcc atggtgttca cggtgctctg cagcctgggc attgtgctgt gcctcgtcac 3300
gtctgtcact gtggagtgga cagggctgaa ggtggccaag agactgcacc gcagcctgct 3360
aaaccggatc atcctagccc ccatgaggtt ttttgagacc acgccccttg ggagcatcct 3420
gaacagattt tcatctgact gtaacaccat cgaccagcac atcccatcca cgctggagtg 3480
cctgagccgc tccaccctgc tctgtgtctc agccctggcc gtcatctcct atgtcacacc 3540
tgtgttcctc gtggccctct tgcccctggc catcgtgtgc tacttcatcc agaagtactt 3600
ccgggtggcg tccagggacc tgcagcagct ggatgacacc acccagcttc cacttctctc 3660
acactttgcc gaaaccgtag aaggactcac caccatccgg gccttcaggt atgaggcccg 3720
gttccagcag aagcttctcg aatacacaga ctccaacaac attgcttccc tcttcctcac 3780
agctgccaac agatggctgg aagtccgaat ggagtacatc ggtgcatgtg tggtgctcat 3840
cgcagcggtg acctccatct ccaactccct gcacagggag ctctctgctg gcctggtggg 3900
cctgggcctt acctacgccc taatggtctc caactacctc aactggatgg tgaggaacct 3960
ggcagacatg gagctccagc tgggggctgt gaagcgcatc catgggctcc tgaaaaccga 4020
ggcagagagc tacgaggggc tcctgggtga gaggctcagg gagaggggtg gggaagagtc 4080
caaggaggag tgtgtctggg ttgggggcca caagggagcc tggggatggg gtggcacttt 4140
cggatactag ctgtggccca tgcctggtgg ctgagcccag cccggccccc agcaccatcg 4200
ctgatcccaa agaactggcc agaccaaggg aagatccaga tccagaacct gagcgtgcgc 4260
tacgacagct ccctgaagcc ggtgctgaag cacgtcatgc cctcatcgcc ctggacagaa 4320
gatcgggatc tgcggccgca ccggcagtgg gaagtcctcc ttctctcttg ccttcttccg 4380
catggtggac acgttcgaag ggcacatcat cattgatggc attgacatcg ccaaactgcc 4440
gctgcacacc ctgcgctcac gcctctccat catcctgcag gaccccgtcc tcttcagcgg 4500
caccatccga tttaacctgg accctgagag gaagtgctca gatagcacac tgtgggaggc 4560
cctggaaatc gcccagctga agctggtggt gaaggcactg ccaggaggcc tcgatgccat 4620
catcacagaa ggcggggaga atttcagcca gggacagagg cagctgttct gcctggcccg 4680
ggccttcgtg aggaagacca gcatcttcat catggacgag gccacggctt ccattgacat 4740
ggccacggaa aacatcctcc aaaaggtggt gatgacagcc ttcgcagacc gcactgtggt 4800
caccatcgcg catcgagtgc acaccatcct gagtgcagac ctggtgatcg tcctgaagcg 4860
gggtgccatc cttgagttcg ataagccaga gaagctgctc agccggaagg acagcgtctt 4920
cgcctccttc gtccgtgcag acaagtgacc tgccagagcc caagtgccat cccacattcg 4980
gaccctgccc ataccccctgc ctgggttttc taactgtaaa tcacttgtaa ataaatagat 5040
ttgattattt cctaaa 5056

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/09797

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(8) : A61K 38/00, 39/00; C07H 21/04; C07K 1/00, 14/00, 16/00; C12N 1/20, 5/00, 15/09; C12P 21/00;
US CL : 424/85.1, 130.1, 184.1; 435/4, 7.1, 69.1, 70.1, 252.3, 320.1, 325; 514/2, 12; 530/300, 350, 387.1

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)
U.S. : Please See Continuation Sheet

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X<br>—<br>A | SHEN et al. Identification of a novel folate receptor, a truncated receptor, and receptor type beta in hematopoietic cells: cDNA cloning, expression, immunoreactivity, and tissue specificity. Biochemisty 1994, Vol. 33, pages 1209-1215, especially Figure 2 and pages 1210-1211 and 1213 (col 2); 95.8% identical to SEQ ID NO: 1 of the instant application. | 1, 3, 6-7, 9, 11, 13, 17, 32, 36-38, 44<br>—<br>2, 4-5, 8, 10, 12, 14-16, 18-31, 33-35, 39-43, 45-173 |

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 01 December 2005 (01.12.2005) | 07 FEB 2006 |
| Name and mailing address of the ISA/US<br>Mail Stop PCT, Attn: ISA/US<br>Commissioner of Patents<br>P.O. Box 1450<br>Alexandria, Virginia 22313-1450<br>Facsimile No. (571) 273-3201 | Authorized officer<br>Bridget E. Bunner<br>Telephone No. (571) 272-1600 |

**Continuation of B.FIELDS SEARCHED Item1:**
424/85.1, 130.1, 184.1; 435/4, 7.1, 69.1, 70.1, 252.3, 320.1, 325; 514/2, 12; 530/300, 350, 387.1; 536/23.1, 23.5; 800/8

G01N 33/53; A01K 67/00

**Continuation of B. FIELDS SEARCHED Item 3:**
EAST, DIALOG (file biosci), MEDLINE, PALM
serach terms: inventors' names, transporter, TRICH